THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 72, NO. 9 NOVEMBER, 1 1947

High Energy Photo-Disintegration of the Deuteron*

M. E. ROSE AND G. GOERTZEL
Clinton Laboratories, Oak Ridge, Tennessee

(Received June 30, 1947)

The disintegration of the deuteron by γ-rays in the range 50-250 Mev is investigated. The lower energy limit is imposed by the neglect of nuclear forces between the neutron and proton after absorption of the photon. Moreover, the effects of retardation and of multipole radiation higher than dipole, the investigation of which constitutes one of the primary purposes of the present work, become apparent only above 50 Mev. The upper limit is imposed by the assumption that the nuclear particles can be treated non-relativistically. It is shown that at the high radiation energies considered it is essential to take into account the range of the nuclear forces in the ground state and that this is the case because of the importance of interference effects (short de Broglie wave-length of the nucleons) which depend very strongly on the "size" of the deuteron. The cross section at 50 Mev is $37\mu b(37\times10^{-30}$ cm$^2)$ and decreases with γ-ray energy up to 150 Mev as $(\hbar\omega)^{-n}$ where n lies between 4 and 5. Beyond 150 Mev the cross section decreases less rapidly because of a more favorable phase relation between outgoing waves from different parts of the deuteron. The general features of the cross section and of the angular distribution of the particles can be understood in terms of interference; the most important effect of the interference is to favor strongly small momentum for the recoil particle (proton in the case of charge coupling and both proton and neutron in the case of spin coupling). The error caused by neglect of nuclear forces in the final state is estimated, by a consideration of electric dipole transitions, to be less than 30 percent in the worst case (low energy photons). The effects of retardation and higher multipoles are calculated explicitly, and it is shown that these effects are small at 50 Mev and important at 100 Mev. It is also shown that the effect of non-central forces between the nucleons in the ground state introduces negligible error.

I. INTRODUCTION

WITH the availability of high energy radiation from the betatron and synchrotron the question of the photo-disintegration of nuclei becomes an interesting one. In general terms the primary problem is to obtain an understanding of the behavior of a nuclear system in the high energy region where the pertinent de Broglie wave-lengths are comparable with the size of the system and, eventually, with the range of the nuclear forces. From the theoretical point of view the case of greatest interest is the dissociation of the deuteron, not only because of the comparative simplicity of this system, but also because only in this case may one expect the photo-particles to be very energetic.

It is therefore our purpose to obtain reasonably reliable values for the photo-disintegration cross section for γ-rays of energies above 50 Mev, where the strong interference effects of high energies begin to appear. Previous investigations[1] which are applicable at low energies involve the neglect of retardation and the neglect of multi-

* This document is based on work performed under Contract No. W-35-058-eng-71 for the Manhattan Project at the Clinton Laboratories.

[1] H. A. Bethe and R. Peierls, Proc. Roy. Soc. **A148**, 146 (1935); E. U. Condon and G. Breit, Phys. Rev. **49**, 904 (1936); H. A. Bethe, Rev. Mod. Phys. **8**, 82 (1936); K. Way, Phys. Rev. **51**, 552 (1937); A Pais, Det. Kgl. Danske Videns. Selskab. Mat.-Fys. Medd. Bind **20**, No. 17 (1943).

poles higher than electric and magnetic dipole.[2] As will be shown below, the retardation and higher multipoles may be neglected for γ-ray energies up to about 50 Mev. For more energetic radiation these effects become significant and must be taken into account. Of even greater importance is the finite range of the nuclear forces. It has already been shown that the finite range has an appreciable effect on the cross section at low energies.[1] At high energies, where the wave-length of the nucleons after absorption of the photon is small, a proper representation of the effects of interference is obtained only by a correct representation of the size of the deuteron, which requires consideration of the finite range of the forces.[3]

An additional motive for investigation of the high energy photo-disintegration might be found in the possibility of obtaining information about nuclear forces in P states or in states of even higher angular momentum. However, because of the small cross sections involved, if for no other reason, the photo-process does not seem eminently suitable for this purpose. In view of the present information about nuclear forces it seemed most appropriate to limit the following considerations to energies such that the effect of the interaction between neutron and proton could be neglected in the final state. For this reason the following calculations were restricted to γ-energies $\geqslant 50$ Mev. An upper limit of 250-Mev photons was imposed by the neglect of relativistic effects.

II. CALCULATION OF THE CROSS SECTION

The differential cross section for the photo-disintegration by a γ-ray of energy $\hbar\omega$, in which the relative motion of the nucleons in the zero-momentum reference frame[4] lies in the solid angle $d\Omega$ is

$$d\sigma=(d\Omega/4\pi)(e^2/\hbar c)(\hbar k/M\omega)(|W|^2)_{Av}. \quad (1)$$

In (1) the symbol Av means an average over the polarization directions of the radiations, over the magnetic substates of the initial state, and a sum over the magnetic substates of the final state. W is the matrix element

$$W=(\Psi_f, w\Psi_i) \quad (2)$$

where the ground state wave function Ψ_i is normalized to unity and the final state wave function is defined so that its asymptotic behavior is a unit amplitude plane wave with the relative momentum $\hbar\mathbf{k}$ plus a converging spherical wave. Where the particles are regarded as free in the final state the spherical wave vanishes and Ψ_f is exactly equal to the unit amplitude plane wave.

In (2) $(e\hbar/Mc)w$ is the perturbation caused by the interaction of the radiation with the charge of the proton and the spins of the two nucleons. We write the vector potential of the radiation as

$$\mathbf{A}(\boldsymbol{\rho})=\boldsymbol{\pi}\exp[i(\mathbf{k}_0\cdot\boldsymbol{\rho}-\omega t)]+\boldsymbol{\pi}^*\exp[-i(\mathbf{k}_0\cdot\boldsymbol{\rho}-\omega t)], \quad (3)$$

where $\boldsymbol{\pi}$ is a unit polarization vector. Then

$$w=i\exp[i\mathbf{k}_0\cdot\mathbf{r}/2]\boldsymbol{\pi}\cdot\text{grad}-\tfrac{1}{2}[\mu_p\boldsymbol{\sigma}_p\cdot\mathbf{H}(\tfrac{1}{2}\mathbf{r})+\mu_n\boldsymbol{\sigma}_n\cdot\mathbf{H}(-\tfrac{1}{2}\mathbf{r})], \quad (4)$$

where $\mathbf{H}$ is the magnetic field derivable from (3), μ_n and μ_p are the magnetic moments of neutron and proton, respectively, in units $e\hbar/2Mc$, and $\mathbf{r}=\mathbf{r}_p-\mathbf{r}_n$ is the vector separation of the particles. The $\boldsymbol{\sigma}_n$ and $\boldsymbol{\sigma}_p$ are the Pauli spin-vectors.

The center of gravity motion has been separated out, giving conservation of total linear momentum as usual, so that the relative coordinates of the particles with respect to the centroid appear as argument in $\mathbf{H}$. Thus (4) has been written in accordance with the fact that because of retardation there is a phase difference in the outgoing waves caused by the spins of the two nucleons.

The first term of (4) gives rise to what will be called the charge transitions, while the remaining terms give the so-called spin transitions.[5] The

[2] The effect of the electric quadripole part of the incident radiation at 17.5 Mev was calculated by W. Rarita and J. Schwinger, Phys. Rev. 59, 556 (1941) and, as expected, the corresponding cross section is small.

[3] The cross section for the photo-disintegration at high energies as calculated by J. M. Jauch, Phys. Rev. 69, 275 (1946) is much too large because of the neglect of the interference arising from the finite range.

[4] That is, the frame in which the total momentum of the system of deuteron plus photon is zero. Thus, before the absorption the deuteron travels in a direction opposite to the photon with energy $E_0=(\hbar\omega)^2/4Mc^2$.

[5] The more customary terminology (photoelectric and photo-magnetic transitions, respectively) seems somewhat ambiguous for the present discussion where the incident

spin operator in (4) can be split into two parts, one symmetric and the other antisymmetric in the spins of the particles. Therefore, while the charge transitions always lead from a triplet ground state to a triplet final state, the spin transitions lead to both triplet and singlet final states. In contrast to the situation at low energies, the former type of spin transition does not vanish. This is due to the retardation and, incidentally, to the neglect of the forces in the final state. Designating the matrix element for the charge transition by W_c and those for the triplet-triplet and triplet-singlet spin transitions by W_{s1} and W_{s0}, respectively, we find

$$(|W|^2)_{Av}=(|W_c|^2)_{Av}+(|W_{s1}|^2)_{Av}+(|W_{s0}|^2)_{Av}. \quad (5)$$

Because the spin transitions involve a change in spin orientation ($\Delta m_s\neq 0$) the charge and spin transitions do not interfere. The two types of spin transition, while non-cohering, are each composed of cohering contributions from the neutron and proton spins.

In the following we shall neglect non-central forces between the nuclear particles in the ground state.[6] Then the total spin is a good quantum number and

$$\Psi_f=\psi_f\chi_{0,1}{}^{m_s} \quad (6)$$

where ψ_f depends on spatial coordinates only and $\chi_{0,1}{}^{m_s}$ is a spin function for singlet and triplet states, respectively. For the ground state

$$\Psi_i=\psi_i\chi_1{}^{m_s'}, \quad (6a)$$

Then

$$W_c=i\boldsymbol{\pi}\cdot(\psi_f,\exp[i\tfrac{1}{2}\mathbf{k}_0\cdot\mathbf{r}]\,\mathrm{grad}\psi_i)\delta_{m_sm_s'}, \quad (7a)$$

$$W_{si}=-\frac{i}{2}(\mathbf{k}_0\times\boldsymbol{\pi})\cdot\mathbf{u}_{m_sm_s'}[\mu_p(\psi_f,\exp[i\tfrac{1}{2}\mathbf{k}_0\cdot\mathbf{r}]\psi_i) +\mu_n(\psi_f,\exp[-i\tfrac{1}{2}\mathbf{k}_0\cdot\mathbf{r}]\psi_i)], \quad (7b)$$

$$W_{s0}=-\frac{i}{2}(\mathbf{k}_0\times\boldsymbol{\pi})\cdot\mathbf{v}_{m_sm_s'}[\mu_p(\psi_f,\exp[i\tfrac{1}{2}\mathbf{k}_0\mathbf{r}]\psi_i) -\mu_n(\psi_f,\exp[-i\tfrac{1}{2}\mathbf{k}_0\cdot\mathbf{r}]\psi_i)], \quad (7c)$$

where

$$\mathbf{u}_{m_sm_s'}=(\chi_1{}^{m_s},\boldsymbol{\sigma}_p\chi_1{}^{m_s'}), \quad (7d)$$

$$\mathbf{v}_{m_sm_s'}=(\chi_0{}^0,\boldsymbol{\sigma}_p\chi_1{}^{m_s'}). \quad (7e)$$

(1) Calculation of the Matrix Elements

For no nuclear forces between particles in the final state, we have simply

$$\psi_f=\exp i(\mathbf{k}\cdot\mathbf{r}), \quad (6a)$$

where the wave vector $\mathbf{k}$ is in the direction of proton momentum in the zero-momentum reference frame. From (7) we obtain

$$(|W_c|^2)_{Av}=\tfrac{1}{2}k^2\sin^2\vartheta I^2(q), \quad (8a)$$

$$(|W_{s1}|^2)_{Av}=(k_0{}^2/12)[\mu_pI(q)+\mu_nI(q')]^2, \quad (8b)$$

$$(|W_{s0}|^2)_{Av}=(k_0{}^2/24)[\mu_pI(q)-\mu_nI(q')]^2, \quad (8c)$$

where ϑ is the angle between the relative motion and the direction of propagation of the photon. The matrix elements are thus seen to depend essentially on the single quantity I defined by

$$I(q)=\int\exp[i\mathbf{q}\cdot\mathbf{r}]\psi_i(r)d\tau \quad (9)$$

and

$$\mathbf{q}=\mathbf{k}-\tfrac{1}{2}\mathbf{k}_0, \quad (10a)$$

$$\mathbf{q}'=\mathbf{k}+\tfrac{1}{2}\mathbf{k}_0. \quad (10b)$$

Evidently $\hbar\mathbf{q}$ and $\hbar\mathbf{q}'$ are the momenta of the neutron and proton, respectively, in the laboratory frame. Therefore, in the case of each interaction term in the hamiltonian (cf. (4)) the transition probabilities depend on the magnitude of the momentum of the recoil particle. Obviously $q'(\vartheta)=q(\pi-\vartheta)$.

To evaluate (9) the ground state wave function for a rectangular potential hole is used. The depth of the hole is V_0 and the range is R. Then with the following notation

$$x=r/R,\quad a=\alpha R=R(M\epsilon)^{\frac{1}{2}}/\hbar,$$
$$b=R[M(V_0-\epsilon)]^{\frac{1}{2}}/\hbar, \quad (11)$$

where ϵ is the binding energy of the deuteron, the wave function ψ_i becomes

$$\psi_i=(2\pi R^3)^{-\frac{1}{2}}(a/1+a)^{\frac{1}{2}}\sin bx/x,\quad x<1,$$
$$\psi_i=(2\pi R^3)^{-\frac{1}{2}}(a/1+a)^{\frac{1}{2}}\sin b\,e^{-a(x-1)}/x,\quad x>1 \quad (12)$$

and $b\cot b=-a$.

---

radiation may be decomposed into the infinite series of electric and magnetic multipoles. Both types of multipoles give rise to transitions caused by coupling with the charge and spins of the particles.

[6] The effect of tensor forces in the ground state is estimated in Section 3 of this paper and is found to be negligible.

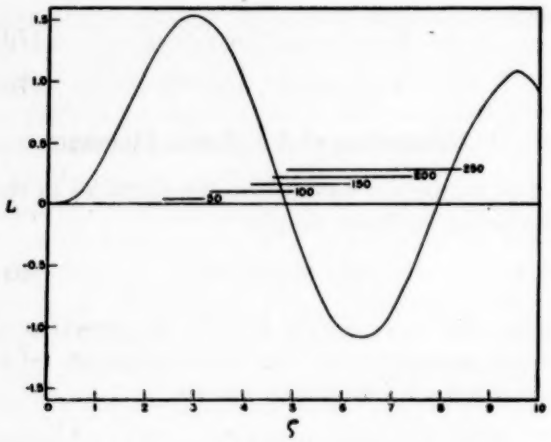


FIG. 1. The quantity L as a function of ζ, the recoil momentum in units $\hbar/R$. The horizontal segments give the range of recoil momenta, corresponding to all possible directions of the photo-particles, for gamma-ray energies (in Mev) given by the affixed numbers.

Introducing the dimensionless "recoil momentum" $\zeta=qR$, we obtain

$$I(q)=(8\pi R^3)^{\frac{1}{2}}(a/1+a)^{\frac{1}{2}}(a^2+b^2) \times \sin b(L(\zeta)/\zeta^4), \quad (13)$$

$$L(\zeta)=\zeta^4/[(a^2+\zeta^2)(b^2-\zeta^2)] \times(\cos\zeta+a\sin\zeta/\zeta). \quad (13a)$$

Because of its importance for the interpretation of the results the quantity $L(\zeta)$ is given in Fig. 1. Obviously the matrix element I decreases quite rapidly with ζ. The values of ζ corresponding to various directions for the emerging particles is given by

$$\zeta^2=R^2(k^2+\tfrac{1}{4}k_0^2-kk_0\cos\vartheta). \quad (14)$$

The horizontal segments in Fig. 1 give the range of ζ corresponding to $0\leqslant\vartheta\leqslant\pi$ for several γ-ray energies. It is clear that for the range of interest the approximation of zero range of the forces gives

$$I(q)=(8\pi\alpha)^{\frac{1}{2}}/(\alpha^2+q^2), \quad (R=0). \quad (13b)$$

This means that the assumption of zero range is an extremely poor one. In this approximation the oscillations in the matrix element as shown in Fig. 1, are entirely overlooked so that the destructive interference between contributions to the outgoing wave arising from different parts of the deuteron is ignored. As a result the cross section is grossly overestimated (see Section 4, below).

(2) Angular Distribution

The angular distribution of the particles in the zero-momentum reference frame $A_0(\mu)$ is obtained from (5), (8), (13), and (14). For the numerical results we have used the generally accepted value $R=2.8\times10^{-13}$ cm, so that $a=0.645$ and $b=1.898$. The results for several γ-ray energies are shown in Fig. 2. In order to facilitate the comparison between different energies all distributions are normalized to the same area.

It is of interest to compare the angular distributions obtained with that to be expected at low energies. At low energies for which only the charge transitions are important the angular distribution is $A_0\sim1-\mu^2$. Since $\mu=1$ corresponds to forward projection of the protons, it is seen that the distribution at 50 Mev corresponds to a definite shift in the direction of forward-going protons. This may be understood as a consequence of the factor ζ^{-8} in the angular distribution which is

$$A_0(\mu)\sim(L/\zeta^4)^2\sin^2\vartheta.$$

and is an expression of the fact that small recoil momenta are always strongly favored in any process arising from a slowly varying interac-

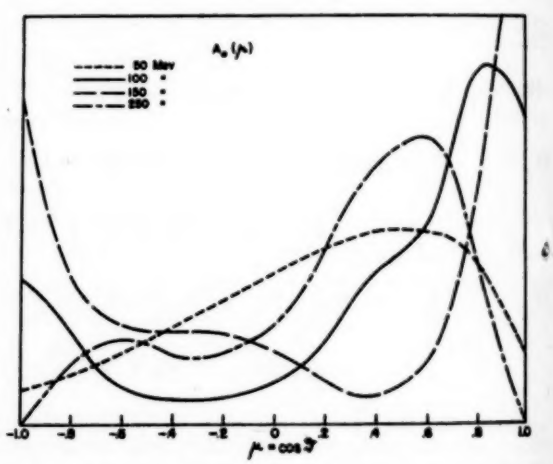


FIG. 2. The angular distribution of the photo-particles in the zero-momentum reference frame. The point $\vartheta=0$ corresponds to protons in the direction of the incident photons. All curves normalized to the same area.

tion.[7] The finite number of particles at $\vartheta=0$ and π is due simply to the spin-transitions which contribute 28 percent to the total number of photo-disintegrations at this energy.

As the energy of the photon is increased to 100 and then to 150 Mev the tendency for forward ejection of the protons becomes more pronounced. This is understood at once by noting that at the higher energies the range of recoil momenta (ζ_{max}/ζ_{min}) is greater and the departure from the $\sin^2\vartheta$ distribution becomes so great that no resemblance to it remains. At the same time it is noted that a peak in the backward direction is developed. This is due to the increasing importance of the spin transitions in which both the proton and neutron play the role of a recoil particle. Superimposed on this distribution with peaks along and opposite to the direction of the photon are characteristic interference minima and maxima which are a reflection of the node in the L function (Fig. 1). This node tends to produce a minimum in the angular distribution at supplementary angles because of the relation between the directions of the recoil momenta of neutron and proton.

At 250 Mev the peaks in distribution are shifted from the extreme positions at $\vartheta=0$ and π and are therefore subdued in magnitude. This is an accidental effect of the destructive interference as may be seen by reference to Fig. 1, which shows that in this case the nodes of L occur near the ends of the angular range.

From the results given and from Fig. 1 it is possible to understand the feature of the angular distribution over the range from 50–250 Mev.

The angular distribution of the protons in the laboratory frame of reference is easily obtained from the foregoing. Designating the angle between the proton motion in the laboratory frame and the direction of propagation of the photon by $\vartheta_0=\arccos\mu_0$ and the distribution function by $A_L(\mu_0)$ we have

$$\mu_0=(\eta+\xi\mu)/(\xi^2+\eta^2+2\xi\eta\mu)^{\frac{1}{2}},$$

[7] At 50 Mev only about 10 percent of the wave-length of the photon is contained within the range of the nuclear forces. Even at 250 Mev the momentum of the photon is not very large. The strong effects of interference described below are mainly due to the fact that the wave-length of the particles, with energies in this range, is so much shorter than the range of the forces.

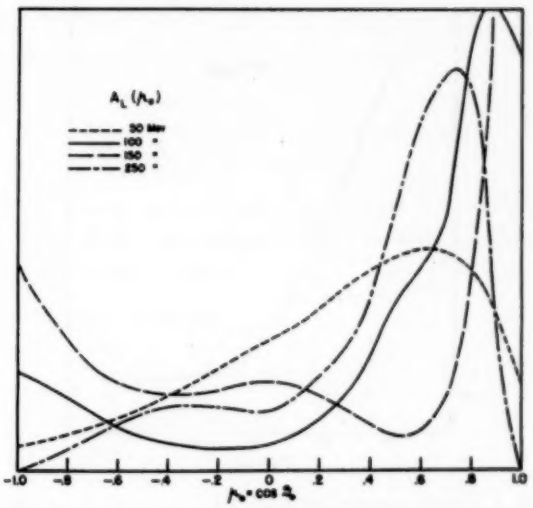


FIG. 3. The angular distribution of the protons in the laboratory frame of reference. The point $\vartheta=0$ corresponds to protons in the direction of the incident photon. All curves normalized to the same area. Ordinate scale same as in Fig. 2.

so that

$$A_L(\mu_0)=A_0(\mu)d\mu/d\mu_0$$
$$=A_0(\mu)(\xi^2+\eta^2+2\xi\eta\mu)^{\frac{1}{2}}/\xi^2(\xi+\eta\mu) \quad (15)$$

where

$$\xi=kR, \quad \eta=\tfrac{1}{2}k_0R. \quad (16)$$

The angular distribution of the protons $A_L(\mu_0)$ is shown in Fig. 3. Again for each γ-ray energy the normalization is the same and the scale in Figs. 2 and 3 is the same. The expected effect of the photon momentum in accentuating the forward peak is clearly evident.

(3) Energy Distribution

Because of the finite momentum of the photon the two nuclear particles do not share the available energy equally. In fact, the angular distribution $A_0(\mu)$ of Fig. 2 also represents the energy distribution of the particles. Designating the energies of proton and neutron, measured in the laboratory frame, by E_p and E_n, respectively, we have

$$E_p=\tfrac{1}{2}(\hbar\omega-\epsilon)+(EE_c)^{\frac{1}{2}}\mu,$$
$$E_n=\tfrac{1}{2}(\hbar\omega-\epsilon)-(EE_c)^{\frac{1}{2}}\mu,$$

where $E_c=(\hbar\omega)^2/4Mc^2$ is the center of gravity energy and $E=\hbar\omega-\epsilon-E_c$ is the kinetic energy of the relative motion in the zero-momentum reference frame. Therefore, Fig. 2 gives the

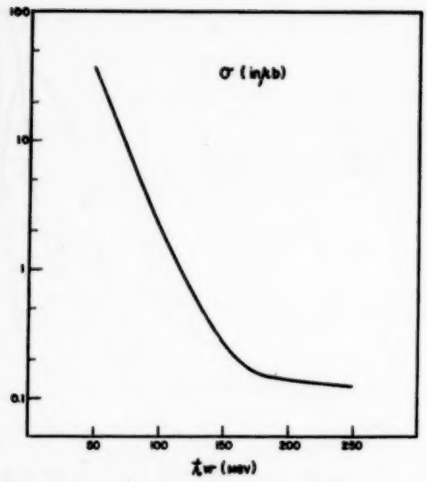


FIG. 4. Total cross section in μb (10^{-30} cm^2) on a logarithmic scale *versus* gamma-ray energy.

energy distribution of the protons (neutrons) if the linear μ-scale is replaced by a linear energy scale extending from the minimum energy $E_{\min}$ at the point $\mu=-1(+1)$ to the maximum energy $E_{\max}$ at $\mu=1(-1)$ and

$$\binom{E_{\max}}{E_{\min}}=\tfrac{1}{2}(h\omega-\epsilon)\pm(EE_c)^{\frac{1}{2}}=\tfrac{1}{2}(E^{\frac{1}{2}}\pm E_c^{\frac{1}{2}})^2.$$

Evidently the proton energy is on the average somewhat greater than the neutron energy because of the charge and the greater magnetic moment of the former.

(4) The Total Cross Section

The total cross section for photo-disintegration is obtained by numerical integration of the angular distribution. The results are shown in Fig. 4 in which the cross section is plotted on a logarithmic scale.

It may be noted first of all that the cross section is rather small in the energy range considered. This arises primarily from the destructive interference caused by the oscillations of the final state wave function inside the potential hole; for example, at 50 Mev, $\xi=kR=3.0$. At this energy the zero-range assumption gives a cross section of $130\mu b$ (μb = micro-barn = 10^{-30} cm^2) to be compared with the present value of $37.4\mu b$. This overestimate by a factor 3.5 is just about what is to be expected from a comparison of the zero range value of L (cf. Eq. (13a)).

$$L=\zeta^4/b^2(a^2+\zeta^2),\quad (R=0) \tag{13c}$$

with the correct value given by Fig. 1. As is to be expected, the situation is aggravated at 100 Mev where the effects of destructive interference are clearly evident. At this energy the cross section found here is $2.16\mu b$ whereas the zero range assumption gives a value about 30 times too large.

The second point to be noted is the rapid decrease in the cross section over the range 50–150 Mev. This is largely due to the fact that the cross section depends sensitively on the minimum recoil momentum which increases with $\hbar\omega$. Thus, the energy dependence of the contribution of the charge transitions is from (1), (9a), and (13)

$$\sigma_c\sim(k^3/\omega kk_0)(\zeta_{\min})^{-8}(L^2\sin^2\vartheta)_{\hbar\omega}$$
$$\sim(L^2\sin^2\vartheta)_{\hbar\omega}/(\hbar\omega)^8,$$

since $\zeta_{\min}=(k-\frac{1}{2}k_0)R\sim kR$. The spin transitions give rise to a contribution σ_s which has an energy dependence given approximately by

$$\sigma_s/\sigma_c\sim k_0^2/k^2\sim\hbar\omega/Mc^2.$$

Considering that the spin transitions become relatively more important with increase of $\hbar\omega$ and noting that $(L^2)_{\hbar\omega}$ decreases with $\hbar\omega$ near 50 Mev and then increases beyond 100 Mev, the energy dependence of the cross section in the region of the rapid decline may be understood. Beyond 150 Mev the cross section falls much more slowly because of the small role played by the destructive interference (see Fig. 1) and to the ever increasing importance of the contribution of the spin transitions.[8]

III. ADDITIONAL CONSIDERATIONS

In the following we present certain considerations bearing on the validity of the assumptions

[8] The small value of the cross section for the photo-effect at high energies explains the fact that the disintegration by high energy electrons is so much more effective than that by photons of the same energy. For example, the electron disintegration cross section at 100 Mev is 180 μb (Bethe and Peierls, reference 1) which is 80 times the cross section for 100-Mev photons. Clearly the electron disintegration is due almost entirely to the soft quanta arising from the contracted Coulomb field. Since the nuclear particles arising from the disintegration under electron bombardment are predominantly of low energy, the assumption of zero range of the nuclear forces should give the correct order of magnitude for the cross section in this case.

on which the foregoing calculations are based. These are (1) the neglect of nuclear forces in the final state and (2) the assumption of central forces in the ground state. In addition, calculations of the effect of retardation and higher multipoles are presented so that the conditions under which these effects can no longer be neglected may be more explicitly stated.

(1) Effect of Nuclear Forces in the Final State

Since the nuclear forces are most effective at the lower γ-ray energies, we consider only those transitions induced by the coupling between the electric dipole part of the radiation and the proton charge. While the neglect of the spin transitions even at 50 Mev is not justifiable, both types of transitions depend on the same matrix element and therefore an estimate of the effect under consideration may be obtained from the consideration of the charge transitions only. The effect of retardation is retained.

With central forces only the final state is ^3P, so that

$$\psi_f = 3ie^{-i\delta}F(kr)\cos\Theta,$$

where δ is the phase shift of the radial function $F(kr)$ and Θ is the angle between $\mathbf{k}$ and $\mathbf{r}$. We represent the P state interaction by a rectangular well of depth V_P and range R. Introducing the notation

$$f_n(y) = (\pi/2y)^{\frac{1}{2}}J_{n+\frac{1}{2}}(y),$$

$$g_n(y) = -(\pi/2y)^{\frac{1}{2}}J_{-n-\frac{1}{2}}(y)$$

where the J's are Bessel functions, we have

$$F = (\cos\delta f_1(\xi) + \sin\delta g_1(\xi))f_1(\xi_P x)/f_1(\xi_P), \qquad x<1 \quad (17a)$$

$$F = \cos\delta f_1(\xi x) + \sin\delta g_1(\xi x), \quad x>1 \quad (17b)$$

where we have used the notation introduced in (11) and (16) and

$$\xi_P = R[M(E-V_P)]^{\frac{1}{2}}/\hbar. \quad (16a)$$

The phase shift is given by

$$\cot\delta = -\frac{\xi_P g_1(\xi)f_0(\xi_P) + \xi g_0(\xi)f_1(\xi_P)}{\xi_P f_1(\xi)f_0(\xi_P) - \xi g_0(\xi)f_1(\xi_P)}. \quad (18)$$

From (7a) we obtain

$$W_c = -(24\pi R/k_0)e^{-i\delta}\sin\vartheta\sin\varphi\delta_{m_s m_s'} \times \int_0^\infty F(\xi x)f_1(\eta x)(d\psi_i/dx)xdx, \quad (19)$$

where ϑ and φ represent the angle between $\mathbf{k}_0$ and $\mathbf{k}$ and the azimuth of the polarization vector π, respectively. Then

$$(|W_c|^2)_{\text{Av}} = (36\pi R/\eta^2) \times \sin^2\vartheta(a^3b/1+a)(Y_i+Y_0)^2, \quad (20)$$

$$Y_i = -\frac{b}{a}\frac{\cos\delta f_1(\xi)+\sin\delta g_1(\xi)}{f_1(\xi_P)} \times \int_0^1 f_1(\xi_P x)f_1(\eta x)f_1(bx)xdx, \quad (20a)$$

$$Y_0 = \frac{\sin b}{b}e^a\int_1^\infty [\cos\delta f_1(\xi x)+\sin\delta g_1(\xi x)] \times f_1(\eta x)e^{-ax}(1+1/ax)dx. \quad (20b)$$

The integrals in Y_i and Y_0 were computed numerically for a series of values of the depth of the P well ranging from 15 Mev attractive to 10 Mev repulsive.[9] The cross section σ_c increases approximately linearly by about $1\mu b$ per Mev repulsion (see Table I). Therefore, the effect of nuclear forces may be expected to change the cross section at 50 Mev by 30 percent or less. At higher energies the correction will, of course, be smaller.

(2) Effect of Retardation and Higher Multipoles

Again restricting our attention to the electric dipole transition, the effect of neglecting retarda-

TABLE I. Cross sections in μb for charge transitions.

Multipoles	V_P(Mev)	Retardation	σ_c(50 Mev)	σ_c(100 Mev)
all	0	with	27.8	0.89
el. dipole	0	with	27.4	
el. dipole	0	without	27.6	0.46
el. dipole	15	with	10.3	
el. dipole	15	without	10.1	
el. dipole	−10	without	39.2	

[9] While no reliable data exist from which the P interaction can be deduced it is of interest to note that the proton-proton scattering at 10 Mev indicates an interaction which certainly falls within the range cited above. See R. R. Wilson, Phys. Rev. 71, 384 (1947).

tion is obtained by replacing $f_1(\eta x)$ in (20a) and (20b) by its value for small argument,[10]

$$f_1(\eta x) \sim \tfrac{1}{3}\eta x.$$

The integrals can then be carried out analytically and we find

$$Y_i = \frac{\eta b}{3a} \frac{\cos\delta f_1(\xi) + \sin\delta g_1(\xi)}{f_1(\xi_P)} \times \frac{b f_1(\xi_P) f_0(b) - \xi_P f_1(b) f_0(\xi_P)}{\xi_P^2 - b^2}, \quad (21a)$$

$$Y_0 = \frac{\eta}{3} \frac{\sin b}{b} \frac{1}{\xi^2 + a^2} \{\cos\delta[a f_1(\xi) + \xi f_0(\xi)(1+1/a)] + \sin\delta[a g_1(\xi) - \xi g_0(\xi)(1+1/a)]\}. \quad (21b)$$

The numerical results for the electric dipole cross section are compared with the corresponding results with retardation and with the cross section due to all multipoles in Table I. The The first two rows of the table show that the effect of higher multipoles is negligible at 50 Mev. The comparison between the second and third as well as between the fourth and fifth rows shows that the effect of retardation is negligible at 50 Mev. On the other hand, from rows one and three it is seen that at 100 Mev the effect of retardation and higher multipoles is important.

(3) Effect of Tensor Forces

The effect of non-central forces in the ground state of the deuteron may be estimated using the model of Rarita and Schwinger.[11] The initial state wave function now becomes

$$\Psi_i = \Phi_{01m} U_S + \Phi_{21m} U_D, \quad (22)$$

where Φ_{LJm} are normalized spin angular functions[11] corresponding to orbital and total angular momentum $L\hbar$ and $J\hbar$, respectively, while $J_z = m$. The radial wave functions for the S and D parts of the ground state are U_S and U_D, respectively. The final state wave function is given by (6) and (6a).

A straightforward calculation gives for the charge transitions

$$(|W_c|^2)_{Av} = 2\pi k^2 \sin^2\vartheta (B_S^2 + B_D^2), \quad (23)$$

where

$$B_S = \int_0^\infty U_S f_0(qr) r^2 dr,$$

$$B_D = \int_0^\infty U_D f_2(qr) r^2 dr.$$

In the absence of tensor coupling $B_D = 0$ and $B_S = B_S^{(0)}$ where

$$B_S^{(0)} = \int_0^\infty U_S^{(0)} f_0(qr) r^2 dr.$$

The radial wave functions are normalized as follows:

$$\int_0^\infty r^2 (U_S^{(0)})^2 dr = 1, \quad \int_0^\infty r^2 (U_S^2 + U_D^2) dr = 1.$$

To a good degree of approximation we can set[11]

$$U_S = \nu_1 U_S^{(0)},$$

where ν_1 is a constant. From the normalization conditions we get[11]

$$\nu_1^2 = 1 - \int_0^\infty U_D^2 r^2 dr = 0.96.$$

Therefore we obtain a lower limit for the charge cross section

$$\sigma_c > \nu_1^2 \sigma_c^{(0)} = 0.96 \sigma_c^{(0)}, \quad (24)$$

where again the superscript zero refers to central forces.

The term B_D^2 which represents the difference between the two sides of the inequality (24) is a small term. In this term we are therefore justified in making the somewhat rougher approximation

$$U_D = \nu_2 U_S,$$

where ν_2 is a constant. This is justified over a

[10] As a check on our results we may reduce the cross section to the Bethe-Peierls result (reference 1) which is readily obtained as the limiting case for zero range and zero P well. Thus, in the limit $Y_i = 0$ and $Y_0 = (\eta\xi/3a) \times (\sin b/b)(\xi^2 + a^2)^{-1}$. With $b = \pi/2$ the cross section becomes (from (1) and (20))

$$\sigma_c = \frac{8\pi}{3} \frac{e^2}{\hbar c} \frac{\hbar^2}{M} \frac{E^{\frac{1}{2}} \epsilon^{\frac{1}{2}}}{(E+\epsilon)^2}.$$

In the limit $R = 0$ the result of II, Section 4 reduces to the cross section given by Jauch (reference 3).

[11] W. Rarita and J. Schwinger, Phys. Rev. **59**, 436 (1941).

region $0.5<r/R<1.5$ from which the integral B_D gets its main contribution. From the normalization we get

$$(1+\nu_2^2)\nu_1^2=1.$$

For small argument qr the function $f_0 \gg f_2$ but for arguments of order π or greater the two functions are of about equal magnitude and differ in sign. Therefore, we replace f_2 by f_0. If anything, this should overstimate B_D. Then

$$B_S^2+B_D^2 \lesssim \nu_1^2(1+\nu_2^2)(B_S^{(0)})^2=(B_S^{(0)})^2 \quad (25)$$

or

$$\sigma_c \lesssim \sigma_c^{(0)}.$$

From (24) and (25) we may conclude that the effect of tensor forces on the charge cross section, and therefore on the total cross section, is 4 percent or less.

Artificial Radioactive Isotopes of Polonium

D. H. TEMPLETON, J. J. HOWLAND, AND I. PERLMAN
Radiation Laboratory and Department of Chemistry, University of California, Berkeley, California

(Received July 18, 1947)

Three new isotopes of polonium have been prepared by bomarding isotope-enriched lead samples with helium ions and bismuth with protons and deuterons. Evidence is given for their assignments as follows:

Isotope	Half-life	Decay	Radiations	Produced by
Po^{206}	9 days	$K(90\%)$, $\alpha(10\%)$	x-rays, e^-, γ0.8 Mev, α5.2 Mev	$Pb^{204}(\alpha,2n)$
Po^{207}	5.7 hours	$K(\sim100\%)$, $\alpha(0.01\%)$	x-rays, e^-(?), γ1.3 Mev, α5.1 Mev	$Pb^{206}(\alpha,3n)$
Po^{208}	~3 years	α	α5.14 Mev	$Pb^{207}(\alpha,3n)$ $Bi^{209}(p,2n)$ $Bi^{209}(d,3n)$

The 9-day Po^{206} has been shown to decay to 6.4-day Bi^{206}. No daughter activities of Po^{207} and Po^{208} have been found. No activity has been found for Po^{209}. Also, 110-min. $F^{18}(\beta^+)$ has been prepared by the reaction $O^{16}(\alpha,pn)F^{18}$ or $O^{16}(\alpha,2n)Ne^{18}\xrightarrow{\text{short}}F^{18}$.

INTRODUCTION

ELEMENTS 81 to 84, thallium, lead, bismuth, and polonium, have many isotopes which are members of the natural radioactive series and whose mass assignments and decay properties are well known.[1] These isotopes all have mass numbers greater than 205 for thallium and 209 for the other three elements. Several additional radioactive isotopes have been prepared by bombardment of the stable isotopes with high energy particles and with neutrons, but knowledge of some of these is much less complete. A thallium activity of 4.23-minute half-life has been assigned to Tl^{206} by Broda and Feather, but details of the work have not yet been published.[2] The mass assignments of two other activities are well established: Pb^{209}, decays by β^--emission with 3.3-hour half-life[3–7] and Tl^{204} decays by β^--emission with 3.5-year half-life.[4, 5, 8–10] Of particular importance to the investigation described in this paper is the isotope of bismuth, which decays by orbital electron capture with 6.4-day half-life and which is prepared by deuteron bombardment of lead[4, 5] and by helium-ion bombardment of thallium.[11] Fajans and Voigt[5] demonstrated that the yield of the isotope depended only upon the Pb^{206} content of the target using 9-Mev deuterons, and therefore that it is Bi^{207} or Bi^{206}, depending on the reaction involved. More recent work[12] indicates that at this energy both the (d,n) and $(d,2n)$ reactions should take place. If 9-Mev deuterons can produce the $(d,2n)$ reaction on Pb^{207} the assignment to Bi^{206} is favored since Fajans and Voigt could find no yield of the 6.4-day activity attributable to the Pb^{207} contents of the targets. Corson, MacKenzie, and Segrè[13] have shown that Bi^{207}

[1] G. T. Seaborg, Rev. Mod. Phys. **16**, 1 (1944).
[2] N. Feather, Phys. Rev. **70**, 88 (1946). Note added in proof: This work has been described recently by E. Broda and N. Feather, Proc. Roy. Soc. London **A190**, 20 (1947).
[3] R. L. Thornton and J. M. Cork, Phys. Rev. **51**, 383 (1937).
[4] R. S. Krishnan and E. A. Nahum, Proc. Camb. Phil. Soc. **36**, 490 (1940).
[5] K. Fajans and A. F. Voigt, Phys. Rev. **60**, 619, 626 (1941).
[6] R. S. Krishnan and E. A. Nahum, Proc. Roy. Soc. (London) **A180**, 321 (1942).
[7] W. Maurer and W. Ramm, Zeits. f. physik **119**, 602 (1942).
[8] M. L. Pool, J. M. Cork, and R. L. Thornton, Phys. Rev. **52**, 239 (1937).
[9] K. Fajans and A. F. Voigt, Phys. Rev. **58**, 177 (1940).
[10] The mass assignment is deduced as 204 or 206 from published observations. The assignment of the 4.23-min. activity to Tl^{206} (reference 2) leaves Tl^{204} for the 3.5-year isotope.
[11] The bombardment of thallium with 40-Mev helium ions will be described in a later paper by the present authors.
[12] For example, see footnote in reference (5)—also, unpublished work of the Manhattan Project.
[13] D. R. Corson, K. R. MacKenzie, and E. Segrè, Phys. Rev. **58**, 672 (1940).

must have a half-life greater than a month by their failure to detect it as the daughter of the alpha-decay of At^{211}.[14] This observation adds further evidence against the Bi^{207} assignment for the 6.4-day activity. Additional evidence for the Bi^{206} assignment is described below.

The existing literature contains the observation and provisional isotopic assignments of several other light isotopes of thallium, lead, and bismuth.[1] Some of these have been encountered in the present series of bombardments and will be reported in detail in later papers together with additional evidence regarding their assignments.

No isotopes of polonium lighter than 210 have been reported. Corson, MacKenzie, and Segrè[15,16] obtained alpha-active material on bombarding lead and bismuth with 32-Mev helium ions but have made a detailed investigation only of the products from bismuth.[13]

These elements present an attractive field for further investigation in this laboratory since beams of 40-Mev helium ions and 20-Mev deuterons are possible with the 60-inch cyclotron.[17] Particles of these energies produce (α,n), $(\alpha,2n)$, $(\alpha,3n)$, (d,p), (d,n), $(d,2n)$, and $(d,3n)$ reactions on these elements in high yield. Most of these reactions produce isotopes of lower mass number than the stable ones, and it is these isotopes about which relatively little is known. High intensity neutron sources now available make possible the production of certain long-lived activities which otherwise could not be detected. In addition, gram quantities of various stable isotopes of lead and thallium have been concentrated in the magnetic isotope separators at this laboratory.[18] Bombardment of different isotopic mixtures permits deductions concerning the reactions taking place and the mass assignments of the products. The results of this program were desired to facilitate the interpretation of the data from bombardments of elements in this region in the 184-inch cyclotron with deuterons of energies up to 200 Mev and helium ions up to 400 Mev, to be described in later papers.

TABLE I. Isotopic compositions of lead samples.

Sample isotope	Natural lead*	Lead *A*	Lead *B*	Lead *C*	Lead *D*
204	1.5%	27.3%	<0.2%	<0.2%	Natural uranium
206	23.6	32.7	59.7	1.9	lead, largely
207	22.6	13.8	25.2	7.8	Pb^{206}, but analy-
208	52.3	26.2	15.1	90.3	sis not available.

* A. O. Nier, J. Am. Chem. Soc. **60**, 1571 (1938).

EXPERIMENTAL

The isotopic compositions of the lead samples used are listed in Table I, as determined by a Nier type mass spectrometer. These samples were purified by recrystallization as lead chloride. A sample of 1 to 50 mg was spread as a slurry of lead hydroxide on a platinum "interceptor" target, dried under an infra-red heat lamp, and heated for 30 minutes at 650°C. The interceptor, of 1-cm^2 area, receives one-third to one-half of the beam of the 60-inch cyclotron. Ordinary lead was bombarded in the form of metal plates large enough to receive the entire beam and mounted on copper disks. Bismuth was bombarded as the oxide mounted on a similar copper disk.

The various samples were bombarded with 20-Mev deuterons or 40-Mev helium ions in the 60-inch cyclotron, usually for 1 to 4 hours. Immediately after bombardment some fraction or all of the lead oxide was dissolved in acetic acid, the lead metal in nitric acid, and the bismuth oxide in hydrochloric acid. Chemical separations were made after addition of inactive carriers when necessary by some combination of the procedures outlined below. The RaE(Bi^{210}) and RaF(Po^{210}) always present were very useful for calculating the chemical yields of the various operations. A measured fraction of the solution before chemical separation was evaporated directly on a platinum counting disk.

Polonium was deposited on a silver disk immersed in a 2-*M* hydrochloric acid solution

[14] Element 85 has been named astatine (symbol At) by D. R. Corson, K. R. MacKenzie, and E. Segrè, Nature **159**, 24 (1947).

[15] D. R. Corson and K. R. MacKenzie, Phys. Rev. **57**, 250 (1940).

[16] D. R. Corson, K. R. MacKenzie, and E. Segrè, Phys. Rev. **57**, 459 (1940).

[17] The bombardments were made possible through the cooperation of Dr. J. G. Hamilton, Mr. T. Putnam, and other members of the group that operates the 60-inch cyclotron.

[18] We are indebted to Dr. B. J. Moyer, Dr. C. M. VanAtta, and members of the isotope separation staff for making available the enriched isotopes, to Dr. E. H. Huffman, Mr. R. C. Lilly, and Miss Dorothy Bockhop for their purification, and to Mr. J. T. Vale for the mass spectrometric assays.

free of nitrate. With vigorous stirring a considerable fraction plated onto the silver in a few minutes, but the yield was not quantitative. Neither bismuth nor lead activities were detected on the silver plates. Another procedure was to heat a dry sample in a small crucible which was covered by a platinum plate on which was set a small beaker of cold water. Some of the polonium can be collected on the platinum in this manner.

Bismuth was precipitated as the iodate from 3-*M* nitric acid solution with an excess of potassium iodate. The precipitate was dissolved in concentrated nitric acid for reprecipitation if desired. The slurry of precipitate was spread on a platinum plate and ignited vigorously in a Bunsen flame to the oxide. This procedure usually removed nearly all of the polonium and resulted in thin, coherent deposits. Another method made use of dithizone (diphenylthiocarbazone) in chloroform which extracts bismuth from aqueous solutions at pH 3.[19] The bismuth was recovered by washing the chloroform with a small volume of dilute nitric acid. Bismuth hydroxide or bismuth iodate was then precipitated and spread on a counting disk.

Lead was precipitated by heating a sulfuric acid solution to fumes of SO_3 and diluting with water. The lead sulfate precipitate was dissolved in ammonium acetate solution. Bismuth impurity was precipitated at this point as the hydroxide. Additional bismuth carrier was sometimes added to improve the bismuth-activity separation. The supernatant was made acidic with acetic acid and the lead was precipitated as lead chromate, which, when spread as a slurry, made uniform counting plates. An alternate procedure proved useful for small quantities. Lead was extracted with dithizone-chloroform solution from a solution which contained oxalate, cyanide, and hydroxylamine at pH 9. Bismuth had to be removed first as described above. Like bismuth, lead was recovered from the chloroform by an acid wash. Thallium could be recovered from the first aqueous residue.

[19] Cf. E. B. Sandell, *Colorimetric Determination of Traces of Metals* (Interscience Publishers, Inc., New York, 1944), pp. 72, 157, 282.

Counting Equipment[20]

Electrons and electromagnetic radiation were measured with argon-ethanol filled Geiger tubes having mica windows of about 3-mg/cm^2 thickness. The geometry factor was about 13 percent for the sample position most frequently used. Electromagnetic radiation, positrons, and negative electrons or beta-particles were differentiated with a similar Geiger tube mounted between the poles of an electromagnet. The sample was mounted in such a way that either the undeflected radiation, the positive particles, or the negative particles were counted. This apparatus also served as a crude magnetic spectrometer for electrons. In some cases a silvered glass-wall Geiger tube, of thickness about 50 mg/cm^2, was also used. Alpha-particles were counted with a Simpson methane proportional counter[21] or in an argon-filled ionization chamber developed in this laboratory. Mixtures of alpha-particles of different energies were analyzed with an alpha-energy analyzer. In this instrument[22] pulses from an argon-filled ionization chamber are amplified and sorted according to size in 48 registering circuits. The pulse resulting from an alpha-particle is recorded on only one of the registers. The ordinal numbers of the registers are a linear function of the amount of ionization produced in the chamber by one alpha-particle and, therefore, over a small range a linear function of the energy of the alpha-particle. A plot of number of alpha pulses recorded against register number is then an alpha particle energy distribution curve (Fig. 3). The energy of an unknown group of particles was estimated by interpolation between alpha-groups of known isotopes which were placed in the ionization chamber at the same time.

Cross Section Measurements

Cross sections were calculated directly only for thin samples mounted on the small inter-

[20] The authors are indebted to Mr. H. P. Robinson, who was responsible for the electronic equipment, and to Mr. R. A. James and Mr. L. O. Morgan who designed the magnetic counter.

[21] J. A. Simpson, Jr., Phys. Rev. **70**, 117 (1946).

[22] A. Ghiorso, A. H. Jaffey, H. P. Robinson, and B. Weissbourd, *An Alpha Pulse Analyzer Apparatus* (Plutonium Project Record, Vol. 14B, 17.3 (1946)) (to be issued).

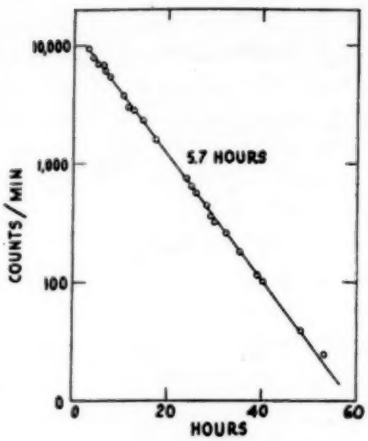


FIG. 1. Decay of gamma-ray of Po^{207}, counted through 0.46 g/cm² aluminum.

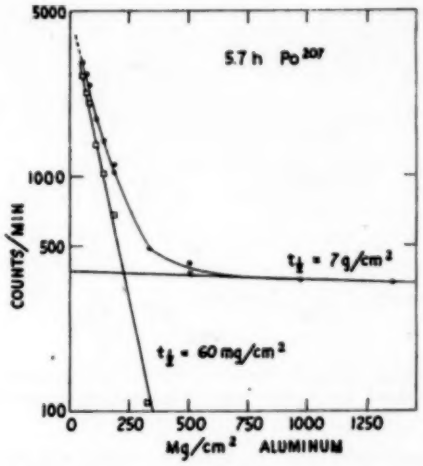


FIG. 2. Aluminum absorption curve for 5.7-hour Po^{207}.

ceptor targets. It was assumed that the beam intensity was not diminished appreciably by the thin sample. The total beam current to the target was integrated by the cyclotron instruments. The number of atoms bombarded was calculated from a gravimetric or colorimetric determination of material in a measured fraction of the solution dissolved after bombardment. The number of atoms transmuted was estimated from counting data for known fractions of this same solution.

RESULTS

140-Day Po^{210}

In every case in which lead was bombarded with 40-Mev helium ions Po^{210}(RaF) was produced by the reaction $Pb^{208}(\alpha,2n)Po^{210}$. Its identity was checked by the alpha-pulse analyzer and by decay measurements. It constituted almost all of the alpha-active material formed when lead *C* (Table I) was bombarded. It was also formed indirectly by decay of 5-day Bi^{210}(RaE) from $Pb^{208}(\alpha,pn)Bi^{210}$, but this contribution was small and was corrected for in all calculations.

Three new activities were observed which followed Po^{210} quantitatively in all chemical separations and which were separated from bismuth, lead, and thallium.

5.7-Hour Po^{207}

Bombardment of lead *C* yielded a 6-hour polonium Geiger activity. It was produced in much better yield from lead *B*, and its half-life was observed as 5.7±0.1 hours for 8 half-lives by counting its gamma-radiation (Fig. 1). An aluminum-absorption curve is shown in Fig. 2. The soft component is probably a mixture of *L* x-rays and conversion electrons. No observations were made on the isolated polonium with a magnetic field, but the original mixture was shown to contain no positron activity of half-life longer than two hours. The gamma-ray has a half-thickness in lead of 12 g/cm², corresponding to 1.3 Mev. These activities are assumed to be due to an orbital electron capture forming an unidentified bismuth isotope.

A weak alpha-activity of half-life about 6 hours was observed in the polonium from lead *B* and lead *D*. In other bombardments it was masked by large yields of other alpha-emitters. With the pulse-analyzer the 6-hour decay was shown to be part of a peak at 5.14 Mev, and not of the Po^{210} peak at 5.30 Mev. Thus the energy of the alpha-particles is about 5.1 Mev.

Yield data presented below for different isotopic mixtures show that the 5.7-hour isotope is made chiefly from Pb^{206} by 40-Mev helium ions. It is not made in appreciable yield from Bi^{209} by 20-Mev deuterons. If the $(\alpha,3n)$ and $(d,3n)$ reactions are prolific while the $(\alpha,4n)$ and $(d,4n)$ are not at these energies, the isotopic assignment must be Po^{207}.

Estimating the $Pb^{206}(\alpha,3n)Po^{207}$ cross section as $1\text{–}2\times10^{-24}$ cm², the counting efficiency of Po^{207} must be 3–5 percent. Since this is higher

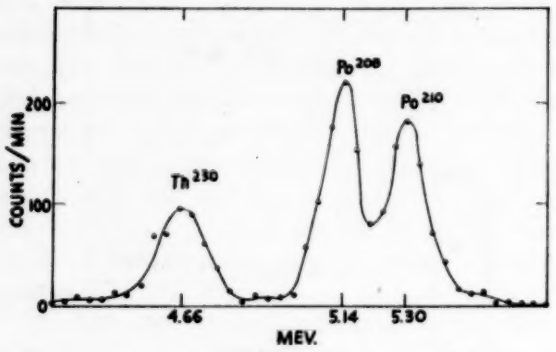


FIG. 3. Pulse analysis curve of Po^{208}, Po^{210} mixture, with standard Th^{230} plate also in ionization chamber.

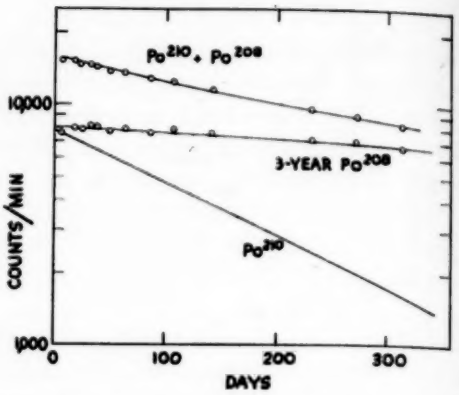


FIG. 4. Decay of Po^{208}, corrected for Po^{210} by means of pulse analysis.

than expected for L x-rays, some of the soft component is attributed to conversion electrons. This counting efficiency corresponds to an alpha-branching ratio of $5\text{–}8\times10^{-5}$, or $\sim$0.01 percent. Thus the half-life for alpha-decay is several years, as expected for an alpha particle of 5.1 Mev in this region.

No daughter of this isotope has been observed. The alpha-branching is so small that observation of its daughter could not be expected in our experiments. The bismuth daughter of the electron capture must be short (less than 15 min.) or very long (many years). In particular, it cannot be the 6.4 day Bi^{206}. From analogy to known isotopes of other odd elements, Bi^{207} is expected to be rather long-lived.

3-Year Po^{208}

Pulse analysis of the long-lived alpha-particles obtained from lead A, lead B, and lead D showed in addition to those of Po^{210} a new group of particles at 5.14 Mev (Fig. 3). Bombardment of bismuth oxide with 20-Mev deuterons produced the same activity in good yield. Lofgren[23] obtained a small amount of long-lived polonium alpha-activity by bombardment of bismuth metal with 15-Mev protons accelerated in the 37-inch frequency-modulated cyclotron. We showed by pulse analysis that it had the same range as this new isotope and that it was free of Po^{210}.

Repeated alpha-pulse analyses for three months showed less than 10 percent decay in the new peak, while that of Po^{210} decayed about 35 percent as expected for its 140-day half-life. Direct decay measurements for ten months, corrected for Po^{210} from pulse-analysis data, showed decay of the new polonium isotope with about a 3-year half-life (Fig. 4). A large sample of this isotope after five months had so little Geiger activity associated with it, if sufficient mica was interposed to stop the alpha-particles, that the half-life of the lead daughter, if it emits x-rays, must be longer than 100 years.

From arguments presented below, based on yields from the different isotopic mixtures, the isotope is assigned to Po^{208}.

9-Day Po^{206}

Bombardment of lead B with 40-Mev helium ions produced a small yield Geiger activity in

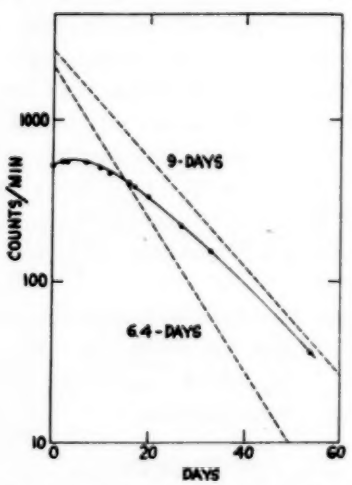


FIG. 5. Growth and decay of Po^{206}, Bi^{206} mixture. The solid line is the theoretical curve obtained by subtraction of the 6.4-day line from the 9-day line.

[23] We wish to thank Dr. E. J. Lofgren for his cooperation in providing us with this sample.

the polonium fraction whose half-life was variously observed over short time intervals as 12 to 20 days. With lead A the yield was much better, and the anomalous decay rate was explained. Figure 5 shows the decay of a sample of polonium purified after the 5.7-hour activity was dead. An initial growth was observed, followed by decay, which suggests that the observed count is due to a parent and daughter of comparable half-life. Bismuth separated from purified polonium a few days old was found to decay with about a 6-day half-life and was concluded to be identical with the well-known 6.4-day Bi^{206}. The observed data are best fitted by assuming the half-life of the parent to be 9 days, and the counting efficiency of the bismuth to be about 1.3 times that of the polonium. If C_0 is the initial counting rate of Po, k the ratio of counting efficiency of Bi to Po, and λ_1 and λ_2 the decay constants of Po and Bi, respectively, the total counting rate at time t for the Po, plus the Bi which has grown in, is:

$$C = C_0[1+k\lambda_2/(\lambda_2-\lambda_1)]e^{-\lambda_1 t} - C_0[k\lambda_2/(\lambda_2-\lambda_1)]e^{-\lambda_2 t}.$$

The two terms of this expression, when plotted on semilog graph paper, are straight lines of slope corresponding to the two half-lives (Fig. 5).

In Fig. 6 is shown an aluminum-absorption curve for the radiations. The soft component has two definite groups of electrons of ranges estimated as approximately 100 and 300 mg/cm² aluminum, corresponding to 400 and 800 kev. The gamma-ray has a half-thickness in lead of 7.8±0.5 g/cm², corresponding to about 800 kev. There is also a weak gamma-component of 1 g/cm² half-thickness (~300 kev). The two electron groups probably are due to partial conversion of the two gamma-rays, the discrepancy in energies being less than the experimental uncertainty. The x-rays make only a small contribution to the observed counting rate, and they have not been resolved in our absorption measurements.

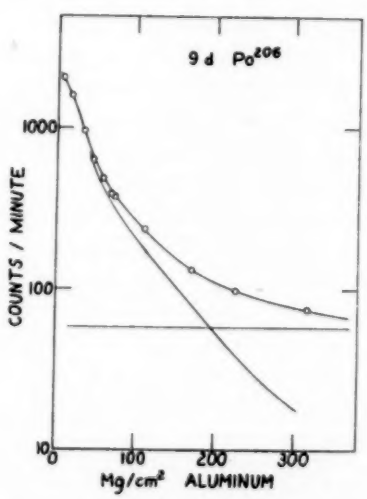


FIG. 6. Aluminum absorption curve for Po^{206}.

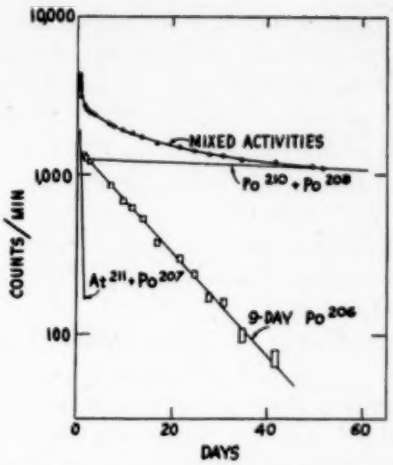


FIG. 7. Alpha-decay curve for polonium from lead A, showing 9-day Po^{206}.

No positrons have been observed in mixtures which contained Po^{206}. Therefore the decay is by orbital electron capture, since a bismuth daughter is produced.

In Fig. 7 is shown the alpha-decay curve for the polonium from lead A. In addition to 7-hour (At^{211} contamination $+Po^{207}$) and >150-day ($Po^{210}+Po^{208}$) periods a component of about 9-day half-life was observed. Pulse analysis failed to resolve this peak from those at 5.14 and 5.30 Mev. The energy is therefore 5.2±0.1 Mev. The yield of this alpha-activity from Pb^{204} was 0.095 that of Po^{210} from Pb^{208} in the same bombardment. If the alpha-activity is due to branching decay of Po^{206} and if the cross sections for $Pb^{204}(\alpha,2n)Po^{206}$ and $Pb^{208}(\alpha,2n)Po^{210}$ are equal, the decay of Po^{206} must be 10 percent alpha and 90 percent electron capture. This corresponds to a counting efficiency of about 18 percent for the gamma-rays, x-rays, and electrons of Po^{206}, and 23 percent for Bi^{206}, for samples mounted on

silver with no correction for back scattering or for window ($\sim$3 mg/cm^2) and air absorption.

If the 9-day alpha-activity is assigned to Po^{205} the yield is too low, and for Po^{204} too high, on the basis of the arguments outlined below concerning $(\alpha,3n)$ and $(\alpha,4n)$ yields. The 90-day half-life for alpha-decay, calculated from the above branching ratio, agrees better with the observed energy of the alpha-particle than the 9-day value does.

The daughter of the alpha-decay has not been observed. It is not the 52-hour lead, whose assignment to Pb^{203} will be discussed in a later paper. If it is Pb^{202} decaying to 13-day Tl^{202}, the half-life of Pb^{202} must be longer than a few weeks. Other experiments also have indicated that Pb^{202} is relatively long-lived.[24]

Yields

In Table II are listed the ratios of the yields of the polonium isotopes from lead *B* to those from lead *C*, together with ratios of isotope concentrations in the two lead samples. From these data it is evident that Po^{210} was made predominantly by the reaction $Pb^{208}(\alpha,2n)Po^{210}$ and that the 5.7-h polonium was made mostly from Pb^{206}. The 3-year polonium was made from Pb^{207} or to some extent from all three lead isotopes. The concentration of Pb^{204} in these samples was so low that it had no effect. The 9-day polonium yield was about 100 times higher from lead *A* than from lead *B* or lead *C*, and therefore this isotope is made from Pb^{204}.

The cross section for the reaction $Pb^{208}(\alpha,2n)Po^{210}$ with 40-Mev helium ions was calculated as 0.4×10^{-24} cm^2.

For the bombardment of bismuth oxide with 20-Mev deuterons, cross sections were calculated as follows:

$$Bi^{209}(d,p)Bi^{210}, \quad \sigma=0.13\times10^{-24}\text{ cm}^2;$$
$$Bi^{209}(d,n)Po^{210}, \quad \sigma=0.036\times10^{-24}\text{ cm}^2;$$
$$Bi^{209}(d,3n)Po^{208}, \quad \sigma=1.1\times10^{-24}\text{ cm}^2.$$

The absolute accuracy is rather uncertain, but the relative values should be correct within a few percent. The ratio $\sigma(d,p)/\sigma(d,n)=3.6$ agrees well with a reasonable extrapolation of the available data at lower energies.[25-28] The absolute values reported by Cork and co-workers[26-28] appear to be inconsistent with this research and with each other.

[24] Bombardment of thallium with 20-Mev deuterons should produce Pb^{202} by the reaction $Tl^{203}(d,3n)Pb^{202}$. In such a bombardment, to be described in a later paper, we failed to observe 13-day Tl^{202}.

TABLE II. Relative yields from lead *B* and lead *C* plus 40-Mev helium ions; and relative concentrations of lead isotopes.

Isotope	Yield *B*/yield *C*	Isotope	%*B*/%*C*
Po^{210}	0.16	Pb^{208}	0.17
3-y. Po	$\sim$3.5	Pb^{207}	3.20
5.7-hr. Po	20.0	Pb^{206}	31.00

Mass Assignments

It has been generally observed for uranium and neighboring elements that the cross sections for a particular type of reaction such as $(\alpha,2n)$ of nuclei with high energy particles are of comparable magnitude for neighboring isotopes.[29] If we assume this generalization for the lead-polonium region, we can deduce the mass assignments of the three new polonium isotopes. The (α,n) reaction is known to be poor for Pb^{207}-$(\alpha,n)Po^{210}$. The cross section for $Pb^{208}(\alpha,2n)Po^{210}$, is about 0.4×10^{-24} cm^2. Leininger and Segrè[30] using the same energy helium ions have shown that for $Bi^{209}(\alpha,3n)At^{210}\xrightarrow{\text{short}}Po^{210}$, the yield is even higher than for $Bi^{209}(\alpha,2n)At^{211}$.

The 9-day polonium is 206 or lighter, since it is made in good yield from Pb^{204}. But its daughter the 6.4-d. Bi, is 206 or 207.[5] Therefore both must be 206, which is confirmed by other arguments stated above for the bismuth isotope.

The low yield of Po^{206} from lead *B* shows that the cross section for the reaction $Pb^{206}(\alpha,4n)Po^{206}$ is very small. The 5.7-hour isotope must be Po^{207} or Po^{208}, since it is made in good yield from Pb^{206}. If Po^{208}, its yield should depend more on the Pb^{207} than it does. The $(d,3n)$ cross section for 20 Mev deuterons is known to be high for heavier elements, and it is expected to be high for bismuth. The very low yield of 5.7-hour

[25] D. G. Hurst, R. Latham, and W. B. Lewis, Proc. Roy. Soc. (London) **174**, 126 (1940).
[26] J. M. Cork, J. Halpern, and H. Tatel, Phys. Rev. **57**, 348, 371 (1940).
[27] J. M. Cork, Phys. Rev. **70**, 563 (1946).
[28] H. E. Tatel and J. M. Cork, Phys. Rev. **71**, 159 (1947).
[29] Unpublished work of the Manhattan Project.
[30] R. F. Leininger and E. Segrè, Private Communication

polonium from bismuth plus deuterons makes the 208 assignment unsatisfactory. Therefore the isotope must be Po^{207}.

The 3-year polonium, made in good yield from Pb^{207}, must be Po^{208} or Po^{209}. The high yield of bismuth plus 20 Mev deuterons confirms this restriction. But the low yield from lead *C* makes the 209 assignment unsatisfactory, for the reaction $Pb^{208}(\alpha,3n)Po^{209}$ should have a high cross section. The observed yield from lead *B* corresponds to $\sigma=0.81\times10^{-24}$ cm². This is in agreement with $Pb^{206}(\alpha,2n)Po^{208}$ ($\sigma=0.4\times10^{-24}$ cm²) and $Pb^{207}(\alpha,3n)Po^{208}$ ($\sigma=3\times10^{-24}$ cm²).

No activity has been observed which can be assigned to Po^{209}. This isotope should be produced in high yield by the reactions $Bi^{209}(d,2n)Po^{209}$ and $Pb^{208}(\alpha,3n)Po^{209}$. If its half-life is short, it must be less than 20 min. If it is long-lived and emits alpha particles of about the same energy as Po^{208}, it would escape detection if its half-life were 10 years or more.

Some of the alpha particles attributed to Po^{208} may have come from Po^{209}, but because of the yield arguments stated above most of them must belong to Po^{208}. If Po^{209} is long-lived and emits x-rays, its half-life must be greater than 1000 years.

110-minute F^{18}

In the various bombardments of lead oxide with 40-Mev helium ions a Geiger activity was observed of half-life 107 ± 5 minutes. About 1 percent of the radiation penetrated 2 g/cm² of lead. With the magnetic counter positions were detected which decayed with a half-life of 115 ±10 min. Chemical separations showed that it was not polonium, bismuth, lead, or thallium. The yield was very erratic when samples were evaporated from acid solution. The activity was carried by a lanthanum fluoride precipitate. This activity is believed to be due to F^{18}, formed by the reaction $O^{16}(\alpha,pn)F^{18}$ or indirectly by $O^{16}(\alpha,2n)Ne^{18}\xrightarrow{\text{short}}F^{18}$. The half-life of F^{18} has been reported variously as 107 to 120 minutes, and it is known to be a positron emitter.[31–39]

7.5-hour At^{211}

In the polonium from lead *A*, two long-range alpha-peaks were observed with the pulse-analyzer. Both decayed with a 7.5 ± 1 hour half-life, and their energies are in agreement within experimental error with the reported values 5.94 Mev[13] and 7.44[40] Mev for At^{211} and Po^{211}, respectively. The yields were in the ratio 44/56 in agreement with the ratio 40/60 reported[13] for At^{211} and Po^{211} in equilibrium. Astatine would have plated on metallic silver and would not have been separated from the polonium.

The observed yield corresponded to an impurity of a few parts per million of bismuth in the lead sample, from which it would be made by the reaction $Bi^{209}(\alpha,2n)At^{211}$. The first component in Fig. 7 is largely due to this activity, and prevented observation of Po^{207} in this case. In other bombardments pulse analysis showed that At^{211} was absent.

This paper is based on work performed under Contract No. W-7405-eng-48 with the Atomic Energy Commission in connection with the Radiation Laboratory, University of California.

[31] A. H. Snell, Phys. Rev. **51**, 143 (1937) (half-life 112 ± 4 min.).
[32] M. L. Pool, J. M. Cork, and R. L. Thornton, Phys. Rev. **52**, 239 (1937) (half-life 108 min.).
[33] L. A. DuBridge, S. W. Barnes, J. H. Buck, and C. V. Strain, Phys. Rev. **53**, 447 (1938) (half-life 107 ± 4 min.).
[34] T. Yasaki and S. Watanabe, Nature **141**, 787 (1938) (half-life 120 min.).
[35] W. L. Davidson, Phys. Rev. **57**, 1086 (1940).
[36] S. B. Welles, Phys. Rev. **59**, 679 (1941) (half-life 114 min.).
[37] R. S. Krishnan, Nature **148**, 407 (1941) (half-life 112 ± 2 min.).
[38] L. B. Borst, Phys. Rev. **61**, 106 (1942).
[39] O. Huber, O. Lienhard, P. Scherrer, and H. Wäffler, Helv. Phys. Acta **16**, 33 (1943).
[40] W. B. Lewis and B. V. Bowden, Proc. Roy. Soc. (London) **A145**, 235 (1934).

Artificial Radioactive Isotopes of Bismuth and Lead

D. H. TEMPLETON, J. J. HOWLAND, AND I. PERLMAN
Radiation Laboratory and Department of Chemistry, University of California, Berkeley, California
(Received July 18, 1947)

Six lead and bismuth activities have been identified from the bombardment of thallium, lead and bismuth with cyclotron deuterons and helium ions, and of lead with pile neutrons.

Isotope	Half-life	Decay	Radiations observed	Produced by
Bi^{204}	12 hours	K	e^-, x-rays, γ	$Tl(\alpha,3n)$ $Pb^{204}(d,2n)$
Bi^{206}	6.4 days	K	e^-, x-rays, γ	Po^{206} K decay $Tl(\alpha,3n)$ $Pb^{206}(d,2n)$ $Pb^{207}(d,3n)$
Bi^{210}	5.0 days	β^-	β^-	$Bi(d,p)$ $Pb^{208}(\alpha,pn)$
Pb^{203}	52 hours	K	e^-, γ	$Tl(d,2n)$ $Pb^{204}(n,2n)$
Pb^{204*}	68 minutes	I.T.	e^-, γ	Bi^{204} K decay $Tl(d,3n)$
Pb^{209}	3.3 hours	β^-	β^-	$Pb(d,p)$ $Pb(n,\gamma)$

No activity has been found which can be attributed to Bi^{205}, Bi^{207}, Pb^{202}, or Pb^{205}, although all of these isotopes must have been produced.

INTRODUCTION

IN this laboratory an extensive study is being made of the light radioactive isotopes of lead and neighboring elements.[1,2] In a previous paper[2] three new polonium isotopes have been described. In this paper some results concerning lead and bismuth isotopes are presented.

EXPERIMENTAL

The experimental techniques which were used have been described in a previous paper.[2] The isotopic compositions of the target materials used are listed in Table I. The enriched samples of Pb^{204} and Pb^{206} were prepared in a calutron in this laboratory.[3]

TABLE I. Isotopic compositions of target materials.

Isotope	Ordinary[a] thallium	Ordinary[b] lead	Lead A[c]	Lead B[c]	Bismuth[a]
Tl^{203}	29.1%				
Tl^{205}	70.9				
Pb^{204}		1.5%	27.3%	<0.2%	
Pb^{206}		23.6	32.7	59.7	
Pb^{207}		22.6	13.8	25.2	
Pb^{208}		52.3	26.2	15.1	
Bi^{209}					100%

[a] A. O. Nier, Phys. Rev. **54**, 275 (1938).
[b] A. O. Nier, J. Am. Chem. Soc. **60**, 1571 (1938).
[c] Analyzed by Mr. J. T. Vale of this laboratory.

These materials were bombarded with 20-Mev deuterons and 40-Mev helium ions in the 60-inch cyclotron.[4] Samples of lead A and lead B were bombarded with a mixture of fast and slow neutrons in the heavy-water moderated uranium pile of the Argonne National Laboratory. Each activity described in this paper was identified as to element by chemical separations.

RESULTS

5.0-day Bi^{210}(RaE)

RaE was produced by the reactions:

$Pb^{208}(\alpha,pn)Bi^{210}$, $\sigma = 0.008\times10^{-24}$ cm^2 at 40 Mev,
$Bi^{209}(d,p)Bi^{210}$, $\sigma = 0.13\times10^{-24}$ cm^2 at 20 Mev.

[1] J. J. Howland, D. H. Templeton, and I. Perlman, Phys. Rev. **71**, 552 (1947).
[2] D. H. Templeton, J. J. Howland, and I. Perlman, Phys. Rev. **72**, 758 (1947).

[3] We are indebted to Dr. B. J. Moyer, Dr. C. M. Van-Atta, and members of the isotope separation staff for making available these materials, to Dr. E. H. Huffman, Mr. R. C. Lilly, and Miss Dorothy Bockhop for their purification, and to Mr. J. T. Vale for the analyses.
[4] The bombardments were made possible through the cooperation of Dr. J. G. Hamilton, Mr. T. Putnam, and other members of the group that operates the 60-inch cyclotron.

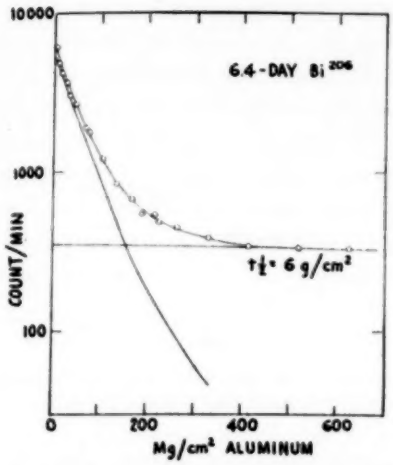


FIG. 1. Aluminum-absorption curve for Bi^{206}.

It was identified by chemical isolation, decay measurement, and observation of the alpha particles of its daughter, Po^{210}. The cross sections are uncertain by 20 percent because of the method used to estimate the intensity of the cyclotron beam. The energy of the incident particles is uncertain by 10 percent, but the targets were thin enough so that there was little degradation of energy.

6.4-day Bi^{206}

The best known light isotope of bismuth is the 6.4-day (orbital electron capture) activity, variously attributed to Bi^{208}, Bi^{207}, or Bi^{206}, and made by bombardment of lead with deuterons.[5,6] In the previous paper[2] its formation from orbital electron capture decay of a 9-day polonium isotope was described, and evidence was given for the assignment to mass 206. It was also formed directly from lead *B* by 40-Mev helium ions. If made by the reaction $Pb^{206}(\alpha,p3n)$, the cross section is of the order of 10^{-26} cm². The presence of a considerable amount of 5.0-day Bi^{210} (from Pb^{208}) made this estimate inaccurate.

Bombardment of metallic thallium with 40-Mev helium ions produced Bi^{206} in good yield by one or both of the reactions $Tl^{205}(\alpha,3n)Bi^{206}$ and $Tl^{203}(\alpha,n)Bi^{206}$. By analogy to lead bombardments,[2] the first of these reactions should predominate.

[5] R. S. Krishnan and E. A. Nahum, Proc. Camb. Phil. Soc. **36**, 490 (1940).
[6] K. Fajans and A. F. Voigt, Phys. Rev. **60**, 619, 626 (1941).

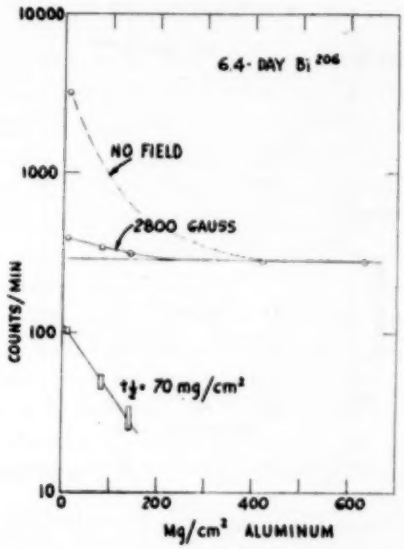


FIG. 2. Aluminum-absorption curve for electromagnetic radiation of Bi^{206}. A magnetic field prevents electrons from being counted.

This activity was also observed whenever lead was bombarded with 20-Mev deuterons.

In Figs. 1 to 4 are shown aluminum- and lead-absorption curves taken with a mica-window brass-wall Geiger counter. In Fig. 2, electrons were deflected with a magnetic field so that the curve represents only the electromagnetic radiations. The energy of the soft component in Fig. 2 is observed as about 14-kev (half thickness 70 mg. Al/cm²) which corresponds approximately to the lead *L* x-rays (9–15 kev). The counting efficiency for Bi^{206} was calculated to be about 20

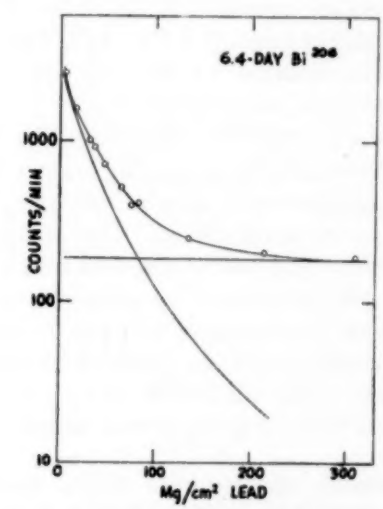


FIG. 3. Lead-absorption curve for Bi^{206}.

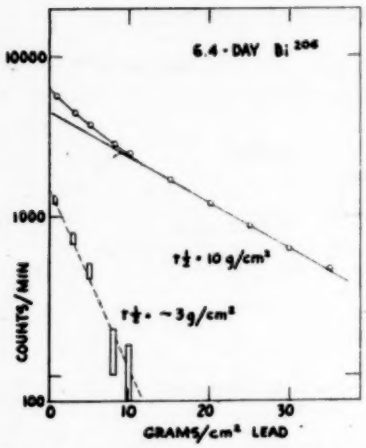


FIG. 4. Lead-absorption curve for Bi^{206} gamma-rays.

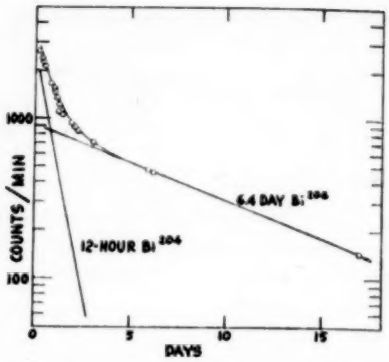


FIG. 5. Decay curve for Bi^{204} and Bi^{206} produced by helium-ion bombardment of thallium.

percent[7] from yield data for Po^{206} in helium ion bombardments.[2] This calculation is based on the assumption that the cross section for $Pb^{204}(\alpha,2n)Po^{206}$ is the same as that for $Pb^{208}(\alpha,2n)Po^{210}$, and on the known counting efficiency and geometry for Po^{210} alpha particles. The relative counting efficiencies of Po^{206} and Bi^{206} are known from the growth and decay curves. From the yield of Bi^{206} so obtained, a 1 percent counting efficiency for the L x-rays is obtained, if there is one per disintegration. The K x-rays were not definitely observed, but the counting procedures were not very sensitive for differentiating these radiations from the gamma-rays.

The lead absorption curve for the gamma-rays, shown in Fig. 4, gives an energy of 1.1 ± 0.1-Mev (half-thickness 10 g Pb/cm²). There appears also to be a softer gamma-ray of about 0.4-Mev energy (half-thickness 3 g Pb/cm²). Fajans and Voigt[6] have reported 1.1 Mev for the gamma-energy based on lead absorption.

Most of the observed counts are due to conversion electrons. The magnetic counter used as a crude spectrometer showed the presence of electrons and no positrons. The spectrum is complex but could not be resolved into its components. The maximum intensity of radiation corresponded to an electron of about 0.3 Mev and one or more weaker lines of higher energy were also present. The aluminum absorption curve for the electrons (Fig. 1) is not simple enough for accurate estimates of the ranges of the individual components. The end-point of about 350 mg/cm² of aluminum (~0.85 Mev) is somewhat higher than the values 0.30 ± 0.02 and 0.2 g/cm² observed by Fajans and Voigt[6] and Krishnan and Nahum,[5] respectively.

[7] This counting efficiency is for samples mounted on silver with no correction for back scattering or for window thickness (3 mg/cm² mica).

12-hour Bi^{204}

An activity of half-life 12 ± 1 hours was observed in the bismuth fraction isolated from thallium bombarded with 40-Mev helium ions (Fig. 5) and lead A (enriched Pb^{204}) with 20-Mev deuterons. A weak activity of 10–16 hour half-life was observed in ordinary lead bombarded with deuterons.

The yield of 12-hour Bi was at least 15-fold greater from lead A than from natural lead. Since the Pb^{204} content of lead A is 18 times that of ordinary lead, while no other lead isotope is similarly enriched, it is concluded that the 12-hour bismuth is formed in appreciable yield only from Pb^{204}. The bombardments were monitored by the yield of Bi^{206}.

The mass assignment of the 12-hour bismuth may be deduced by the same sort of argument as those used in the previous paper[2] for polonium isotopes. Since 20 Mev deuterons produce the activity from Pb^{204} in good yield, it must be Bi^{203}, Bi^{204}, or Bi^{205}. Since 40 Mev helium ions produce it from thallium, it is probably Bi^{204} or heavier. Since it is not produced in good yield from Pb^{206}, Pb^{207}, or Pb^{208} with 20-Mev deuterons, it cannot be Bi^{205} or heavier. Therefore it has been assigned Bi^{204}.

Since the cross section for $Tl^{203}(\alpha,4n)$ might possibly be high enough to account for the ob-

served yield of the 12 h Bi, the assignment Bi^{203} is a remote possibility. However, the assignment of the 52 h Pb activity to Pb^{203} is quite definite, and this activity is not a daughter of the 12 h Bi.

Krishnan and Nahum[5] have reported a very weak 18±2-hour activity in the bismuth from lead bombarded with 9 Mev deuterons. We have not observed this activity unless it is the same as this 12-hour isotope.

In Fig. 6 is shown an aluminum absorption curve for a mixture of the 12-hour and 6.4-day bismuth activities. Subtraction of the contribution due to Bi^{206} calculated from the curves in Figs. 1 and 5 gives a curve for the 12-hour activity alone. About 7 percent of the counts are due to the gamma-rays, but there are insufficient data for an estimate of the energy.

The magnetic counter showed that only about 10 percent of the counts are due to electromagnetic radiations. About 7 percent are due to the gamma-rays, and so, like Bi^{206}, very little of the soft component is due to L x-rays. Examination of the charged particles with the magnetic counter showed negative electrons of energy about 200 kev, and a smaller number with energies as high as 750 kev. The latter value is uncertain because of the presence of a considerable amount of Bi^{206}. No positrons were observed. The absorption curve for these particles (Fig. 6) has a soft component of range 30 to 50 mg/cm² Al, (~0.2 Mev) and a harder component of range 300 mg/cm² Al or more (~0.8 Mev).

When bismuth containing the 12-hour activity was purified, allowed to stand, and then separated from added inactive lead, the lead was found to contain the 68-minute period attributed to Pb^{204*}.

In a series of such lead separations from aliquots of a bismuth fraction, the yield of 68-minute lead fell off with an 11±2 hours half-life. Thus 68-minute lead is undoubtedly the daughter of the 12-hour bismuth by orbital-electron capture. The yield of lead was about 4 counts for each 100 counts of 12-hour bismuth present. Because of the low yield, the growth of lead daughter is not apparent on the bismuth-decay curves. The 68-minute lead activity is discussed in detail below.

68-minute Pb^{204*}

A one-hour lead activity has been produced by bombardment of lead with fast neutrons (68-minutes,[8] 80-minutes[9]), and thallium with deuterons (65±5 minutes[6]). Fajans and Voigt[6] reported the yield from thallium with 9-Mev deuterons as 1/150 that of 52-hour lead. Krishnan and Nahum[5] did not find the activity in similar deuteron bombardments.

It was not found in 9-Mev deuteron bombardment of lead.[6] Using ordinary lead, uranium lead, and thorium lead, Maurer and Ramm[8] showed that the activity is made from Pb^{204} by neutrons of 0.8-Mev and higher energy, and that the cross section with thermal neutrons is small. The cross section for formation from Pb^{206} is very small. From these facts they deduced that the activity must be either Pb^{204*} or Pb^{205}. The low yields from thallium plus deuterons, from lead plus thermal neutrons, from lead plus deuterons, and from Pb^{206} plus fast neutrons were cited as evidence against the Pb^{205} assignment.

In the present work the activity was produced from thallium with 20 Mev deuterons, probably by both the reactions $Tl^{203}(d,n)Pb^{204*}$ and $Tl^{205}(d,3n)Pb^{204*}$.

Helium ion bombardment (40 Mev) of thallium produced the activity as the daughter of Bi^{204}, formed by $Tl^{203}(\alpha,3n)Bi^{204}$, There is no evidence for its direct formation by $Tl^{203}(\alpha,p2n)$-

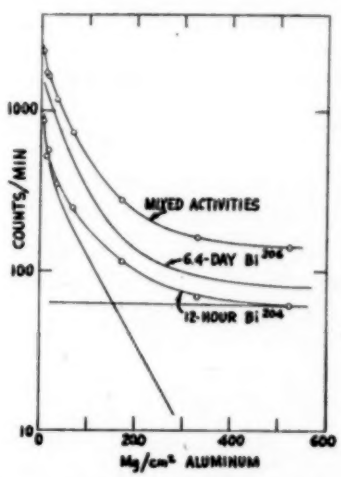


FIG. 6. Aluminum-absorption curve for Bi^{204} and Bi^{206} mixture.

[8] W. Maurer and W. Ramm, Zeits. f. Physik **119**, 602 (1942).
[9] H. DeVries and G. Diemer, Physica **6**, 599 (1939).

Pb^{204*}. The activity was observed in 20-Mev deuteron bombardment of enriched Pb^{204} by the mechanism $Pb^{204}(d,2n)Bi^{204}\xrightarrow{12h\ K}Pb^{204*}$. There is no evidence of its direct production by $Pb^{204}(d,d)Pb^{204*}$.

Our pile-neutron bombarded lead was examined too late for detection of a 68-minute period.

The properties of the radiations have been described by previous workers.[6,8] We have observed gamma-rays, negative electrons of several hundred-kev energy, and no positrons, in accord with the published data. With our counters, about 7 percent of the counts are due to gamma-rays, which is about the same as the ratio for Bi^{204}. Because of this fact it seems improbable that the counting efficiencies of the two activities are very much different, and so only about 4 percent of the disintegrations of Bi^{204} produce 68-minute Pb^{204*}. The other disintegrations presumably give stable Pb^{204}.

3.3-hour Pb^{209}

There is general agreement that a 3.3-hour lead β^--emitter produced by deuteron bombardment of lead[5,6,10] and by neutron bombardment of lead and bismuth[8] is due to Pb^{209}. Maurer and Ramm[8] have summarized the properties of the radiations and the arguments for the mass assignment.

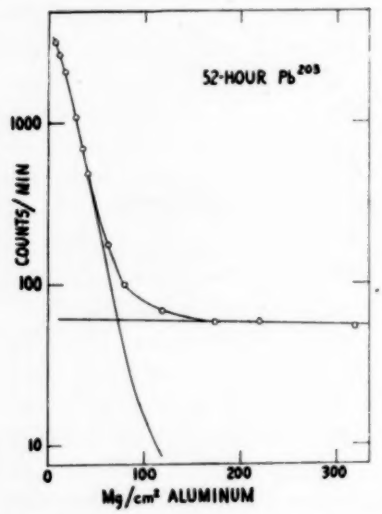


FIG. 7. Aluminum-absorption curve for Pb^{203}.

[10] R. L. Thornton and J. M. Cork, Phys. Rev. **51**, 383 (1937).

In the bombardments of the various lead samples with 40-Mev helium ions, the lead fractions showed a weak 3.3-hour activity mixed with other periods, all of which could be explained as contamination by the bismuth and polonium activities which are formed in much higher yield. The amount of Pb^{209} could be accounted for by a reasonable estimate of the neutron flux in the neighborhood of the target, using 10^{-27} cm^2 for the capture cross section of Pb^{208} for thermal neutrons.[8] There is no evidence for the $Pb^{208}(\alpha,n2p)Pb^{209}$ reaction, but if it occurs the cross section is less than 10^{-29} cm^2.

The 3.3-hour activity was also observed, as expected, in all 20-Mev deuteron bombardments of lead. In the pile-neutrom bombardments, the samples were not received until after this activity had decayed too much for observation.

52-hour Pb^{203}

A 52-hour lead activity has been prepared by deuteron bombardment of thallium[5,6,11,12] and by neutron bombardment of lead.[5,8] Maurer and Ramm[8] bombarded ordinary lead, thorium lead, and uranium lead with 5.3-Mev and 15.2-Mev neutrons. The 52-hour activity was produced only with ordinary lead and 15.2-Mev neutrons, and they concluded it to be formed by the reaction $Pb^{204}(n,2n)Pb^{203}$. Krishnan and Nahum[5] reported the formation of the activity with both "fast" and "slow" neutrons. The activity was not observed in 9-Mev deuteron bombardments of lead.[5,6] The assignment has also been suggested as Pb^{204*}, Pb^{205}, and Pb^{206*}.

We did not observe this activity in any 20-Mev deuteron bombardment of lead, including bombardment of enriched Pb^{204}. This fact makes the Pb^{205} assignment very unlikely. It was observed in lead *A* (27.3 percent Pb^{204}) bombarded with pile neutrons but not in lead *B* (<0.2 percent Pb^{204}), confirming the conclusion of Maurer and Ramm[8] that it is made only from Pb^{204}. The samples were placed in the pile at a point where both fast and slow neutrons were present. This activity was also produced in good yield by bombardment of thallium with 20-Mev deuterons, by the reaction $Tl^{203}(d,2n)Pb^{203}$. These experi-

[11] K. Fajans and A. F. Voigt, Phys. Rev. **58**, 177 (1940).
[12] R. S. Krishnan and E. A. Nahum, Proc. Roy. Soc. (London), **A180**, 321 (1942).

ments confirm the assignment of this activity to Pb^{203}.

The absence of the activity in lead plus 20-Mev deuteron bombardments indicates that Bi^{203} is moderately long-lived, because it should have been produced, and it almost certainly decays by electron capture to Pb^{203}.

An aluminum-absorption curve for this activity (Fig. 7) agrees very well with that published by Maurer and Ramm.[8] The soft component was shown to be mostly a line of negative electrons of energy ~150 kev by use of the magnetic counter. The resolution of this counter is not good enough to show the fine structure reported for these electrons.[13]

The published values[5,6,8] for the maximum electron energies from aluminum absorption vary from 330 to 500 kev, but our data with the magnetic counter indicate a very low abundance of such harder electrons. Possibly some of this harder radiation is due to x-rays.

Our lead-absorption curves are in agreement with published data[6,8] but are not extensive enough for resolution of the gamma-ray components.

Other bismuth isotopes

We have found no activities corresponding to Bi^{205} or Bi^{207} which should be formed from lead by 20-Mev deuterons. These isotopes are expected to be longer lived than the adjacent even-mass isotopes, and they may be very much longer.

Pb^{205} and Pb^{202}

In the bombardment of thallium with 20-Mev deuterons both Pb^{205} and Pb^{202} should have been produced in good yield. No activities were found which could be attributed to these isotopes. Pb^{202} should decay by electron capture to 13-day Tl^{202},[5,6,8] which was not observed in the bombardment. Therefore Pb^{202} is long-lived. One expects that Pb^{205} also is long-lived. If either isotope has a counting efficiency the same as Pb^{203}, its half-life must be greater than 500 years.

Other lead activities

Several other periods have been reported as belonging to lead isotopes, but none of these has yet been confirmed by independent workers. None of these activities was specifically looked for in the experiments reported here.

Krishnan and Nahum[5] reported an intense 10.25±0.25-minute positron activity in lead separated from thallium which had been bombarded with 9-Mev deuterons. The activity was not produced with 7-Mev deuterons. Fajans and Voigt[6] made several attempts to find this activity by the same method without success. Maurer and Ramm[8] could not produce it with fast neutrons on lead.

Pool, Cork and Thornton[14] have reported weak 5-minute and 1.5-hour activities in lead bombarded with fast neutrons. DeVries and Diemer[9] concluded by chemical separations that the former period was not lead or thallium, but some impurity, and found 80-minutes for the latter. Bretscher and Cook[15] later reported the formation of a 4.6-minute thallium activity by fast neutron bombardment of lead. Probably in both cases the longer period is a mixture of 68-minute and 3.3-hour lead.

Waldman and Collins[16] have reported a very weak 1.6±0.2-minute activity produced from lead by x-rays of energy greater than 0.65 Mev. It was not found by Maurer and Ramm[8] using fast neutrons on lead.

This paper is based on work performed under Contract No. W-7405-eng-48 with the Atomic Energy Commission in connection with the Radiation Laboratory, University of California.

[13] A. L. Lutz, M. L. Pool, and J. D. Kurbatov, Phys. Rev. **65**, 61 (1944).

[14] M. L. Pool, J. M. Cork, and R. L. Thornton, Phys. Rev. **52**, 239 (1937).

[15] E. Bretscher and L. G. Cook, Nature **146**, 430 (1940).

[16] B. Waldman and G. B. Collins, Phys. Rev. **57**, 338 (1940).

The Radioactive Decay of Slow Positive and Negative Mesons*

G. E. VALLEY
Laboratory of Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received July 21, 1947)

Cosmic ray mesons, after passage through a cloud chamber situated in a magnetic field, were slowed down and allowed to decay in specimens of carbon, stainless steel, brass, water and beryllium. By means of delayed coincidence circuits the cloud chamber was actuated only when a meson decayed in the specimen. In stainless steel and in brass, only positive mesons decay; in carbon, water, and beryllium both positive and negative mesons decay in such numbers as to indicate that few, if any, of the negative mesons undergo nuclear capture in these substances. It is shown that the masses of the decaying mesons are approximately two hundred times the mass of an electron.

I. INTRODUCTION

IT has been established that a cosmic ray meson at rest can decay radioactivity into at least one high-speed charged particle, and that its half-life for radioactive decay is about 2 μsec. These results have been confirmed by a considerable number of experimenters all using essentially the same "delayed coincidence" technique.[1–7]

Conversi, Pancini, and Piccioni[8] have reported such an experiment in which the delay-coincidence Geiger counters were surmounted by an analyzing magnet so that alternatively only positive or only negative mesons could be brought to rest within range of the counters. They found that when the mesons were brought to rest in iron, only the positive ones decayed radioactively, whereas when brought to rest in carbon (graphite), mesons of both signs decayed radioactively.

This finding has excited considerable attention among theorists[9–12] since it seems to contradict predictions of the meson theories of nuclear forces. According to these theories, the meson, once within range of the nuclear binding force, should be absorbed in a time so short compared to its natural half-life that any particles emitted in the process could not be detected by the delayed coincidence method. Since, however, the electric field of the nucleus keeps the positive mesons from coming within range of the nuclear binding force, they alone should be able to decay radioactively. Thus, the findings of Conversi *et al.*, for the case of mesons stopped in iron, should be typical of all elements which might be used to stop mesons.

Sigurgiersson and Yamakawa[13] have measured the absolute rate of delayed coincidences in various elements and find values consistent with the results of Conversi *et al.*

The experiment described here was already begun when the results of Conversi *et al.* were made known in this laboratory. Their results have been confirmed and extended to other elements.

II. EXPERIMENTAL METHOD

Arrangement of Apparatus

A schematic diagram of the apparatus is shown in Fig. 1. Three G.M. counter trays labeled 1, 2, and 3 contained respectively four, three, and three counter tubes effectively in parallel, and

* A report of the results for carbon and iron was made to an informal group of members of the A.P.S. in Washington, D. C. on May 1, 1947.

[1] F. Rasetti, Phys. Rev. **59**, 613 (1941); **60**, 198 (1941).
[2] P. Auger, R. Maze, and R. Chaminade, Comptes rendus, **213**, 381 (1941).
[3] R. Maze and R. Chaminade, Comptes rendus **214**, 266 (1941).
[4] B. Rossi and N. Nereson, Phys. Rev. **62**, 417 (1942); **64**, 199 (1943).
[5] R. Maze, R. Chaminade, and A. Freon, J. de phys. et rad. **7**, 202 (1945).
[6] M. Conversi and O. Piccioni, Phys. Rev. **70**, 859, 874 (1946).
[7] H. K. Ticho, Phys. Rev. **71**, 463 (1947).
[8] M. Conversi, E. Pancini and O. Piccioni, Phys. Rev. **71**, 209 (1947).
[9] J. A. Wheeler, Phys. Rev. **71**, 462 (1947); **71**, 320 (1947).
[10] E. Fermi, E. Teller, and V. Weisskopf, Phys. Rev. **71**, 314 (1947).
[11] G. Gamow, Phys. Rev. **71**, 550 (1947).
[12] V. Weisskopf, Phys. Rev. **72**, 155 (1947).
[13] T. Sigurgiersson and A. Yamakawa, Phys. Rev. **71**, 319 (1947).

were connected in threefold instantaneous coincidence. They defined a beam of particles to pass through the cloud chamber *CC* and the magnetic field *H*. After passage through tray 3, these particles penetrated into specimen block *S* ann a few of them stopped therein. Surrounding specimen block *S* were sixteen counters connected effectively in parallel, and together comprising tray *D*. Plates L_1 and L_2 were of lead 1.5 and 0.5-inches thick respectively. They served to eliminate most of the soft component of the cosmic radiation and so lowered the counting rates. Attention is called to water jackets J_1 and J_2, each of which was equivalent to about $\frac{1}{8}$ inch of brass. One of these, J_2 had a useful effect upon the measured momentum spectrum as will be described below. All of the counters were electrically shielded; in cases where this shielding was of such a nature as to affect the experimental results, it is indicated as a dashed line in Fig. 1.

Whereas Conversi *et al.*[8] allowed alternatively only positive or only negative mesons to fall on the specimen material and then measured the resulting delayed coincidence rate, in the present experiment a different procedure was followed. All particles of either sign which could penetrate to it, were allowed to fall upon the specimen block *S*, the magnetic field *H* causing no discrimination against either positive or negative particles. Then, whenever tray *D* was discharged in an interval of from one to approximately ten microseconds after a coincidental discharge of trays 1, 2, and 3, the cloud chamber was tripped, and the curved track of the causitive particle photographed. Thus if negative tracks were observed it meant that negative mesons decayed radioactively. The half-life of the mesons was not measured.

Electronic Details

The counters of trays 1, 2, 3 and of the vertically arrayed portion of tray *D* had each an active length of ten inches. The seven counters forming the horizontal array of tray *D* had each an active length of twenty inches. All the counters were constructed of brass tubing of one-inch outside diameter with a wall thickness of 0.032 inch. The tubes were filled with the usual mixture of argon and ethyl alcohol vapor required for self-quenching.

The circuits employed were remarkable only for some special characteristics not of interest to this experiment. They are shown in block diagram form in Fig. 2.

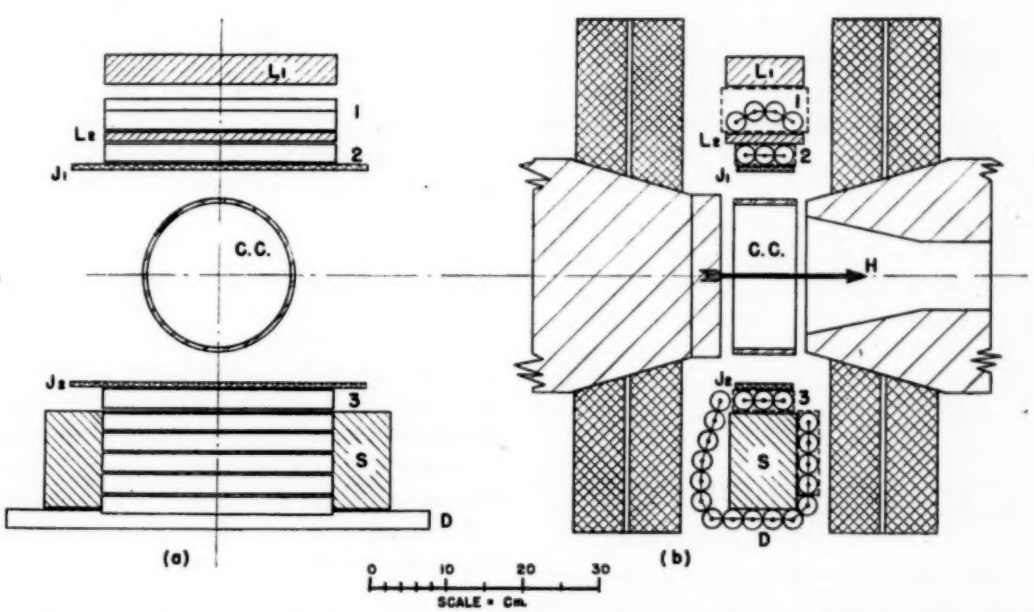


FIG. 1. Disposition of apparatus: 1, 2, 3 designate counters of the meson beam defining "telescope"; *D* designates the counters which surrounded the specimen and which responded to the product particle of the decaying meson; *S* is the specimen; L_1 and L_2 are lead slabs; J_1 and J_2 are brass-water-jackets. *CC* is the cloud chamber and the arrow *H* designates the magnetic field.

Each counter tray shown in Fig. 1 was divided into subgroups of three tubes or less, connected to common high voltage charging resistors. The members of each subgroup were chosen to have equal threshold voltages. The operating voltages of the several subgroups were separately adjusted so that all gave pulses of approximately equal amplitude. Each subgroup was connected to its own type 6AK5 pentode, operating at normal plate voltage, and upon firing, drove that tube's grid to a potential much below cut-off. All the type 6AK5 tubes of a given counter tray shared a common plate resistor, and together with an output cathode follower (type 6J6) comprised one of the preamplifiers *a*, *b*, *c*, and *d* (Fig. 2). The output pulses from these amplifiers rose to their maximum amplitude in 0.2 μsec. and appeared across an output impedance of a few hundred ohms. In effect, all the tubes of a particular tray functioned as one tube, because of the signal addition property of the common anode resistor circuit when the dynamic plate resistances are large. The charging resistors were of 2 megohms, the capacity of an individual counter tube was about 10 μμf, of its coupling condenser 50 μμf. In some cases, the restricted space of the apparatus made it necessary to have about 15 μμf capacity to ground due to cabling. The mean counter overvoltage was 125 volts.

The pulses from preamplifiers *a*, *b*, and *c* were applied to a triple coincidence circuit (*IIIC* in Fig. 2) whose resolving time was 1 μsec. In parallel with each input terminal of this coincidence circuit was connected an individual scale of four, whose function was to monitor the counting rates of the individual trays. Similarly and for the same reason, the output pulse of tray *D*, besides being applied to the delay-coincidence set (*DD* in Fig. 2) was applied also to a scale of thirty-two.

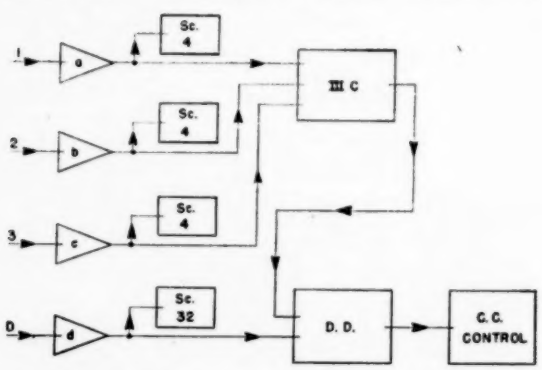


FIG. 2. Block diagram of circuits used; *a*, *b*, *c*, and *d* are preamplifiers used for counters 1, 2, 3 and *D* respectively; *IIIc* is a three-fold coincidence circuit; *DD* is a delayed-coincidence circuit which responded whenever it received a pulse from *d* within the interval from 1 to 10 μsec. after it had received an initiating pulse from *IIIc*; "cc control" is the cloud-chamber-control circuit.

The delayed-coincidence set** functioned in the following way: The pulse from the triple-coincidence set upon entering it was sharpened by a biased-off blocking oscillator circuit, delayed for 1 μsec. by means of a delay line, and then caused to initiate a positive pulse of 9 μsec. duration. This long pulse was applied to the third grid of a type 6AS6 pentode, to whose first grid was applied the positive pulse from pre-amplifier *d*, also after sharpening. The type 6A6S tube is designed to have short grid-bases for both No. 1 and No. 3 grids, and it functions very well as a double-coincidence element. Its plate current cannot flow unless both grids are simultaneously more positive than their cut-off voltages. The signal which appeared at the plate of this pentode was used to trip an electro-mechanical register and also the cloud-chamber control circuit (*CC* in Fig. 2).

By means of an auxiliary circuit in the delayed-coincidence set, the pulses from pre-amplifier *d* and from the threefold coincidence set could be applied directly to a simple-coincidence element. The resultant pulses marked the simple fourfold coincidences between trays 1, 2, 3 and *D*. Besides being continuously registered, these could also be used to trip the cloud chamber.

The Cloud Chamber and Associated Equipment

The cloud chamber was of the rubber diaphragm type filled with argon and the usual 70–30 solution of ethyl alcohol and water. Its illuminated volume was 7 inches in diameter and 2¼ inches deep. It operated at a mean absolute pressure of 93 cm of mercury.

The magnetic field was supplied by a cast-steel yoke*** one pole of which was hollow so

** Mr. M. L. Sands kindly loaned this piece of equipment, which he had originally constructed to measure the spontaneous delays of counter tubes.

*** This was patterned after a design of Professor R. B. Brode.

that photographs could be taken through it. The magnet was energized by an unregulated generator and over the several months' course of the experiment its field varied at most by plus or minus 200 Gauss around a mean value of 5800 at the center of the active volume of the cloud chamber. Over most of the volume of the cloud chamber, the field was uniform to ±1 percent; near the hollow pole, however, the field rapidly became non-uniform, at its worst being non-uniform by ±5 percent.

The use of an electromagnet in conjunction with a vertical cloud chamber for cosmic-ray studies has been noted to cause the tracks to be seriously distorted because of convection currents existing in the chamber prior to its expansion. These are due to unequal heating of the chamber.[14]

This problem arose during the present work. It was overcome by surrounding the cloud chamber with jackets through which water at controlled temperature was circulated. Six thermopiles were fastened to the cloud chamber at various points and by means of a recording potentiometer indicated temperatures to ±0.1°C. By judicious regulation of the water system, these six thermopiles could all be kept within ±¼°C of the same temperature.

Scattering

The effect of coulomb scattering upon the measured curvature was calculated. From Eq. (153a) of Rossi and Greisen's paper,[15] the root-mean-square angular deflection of a particle of given velocity traversing 10 cm. in argon at 1.2 atmospheres pressure was calculated. This was divided by the angular deviation produced on the same particle by a field of 5800 Gauss over a 10-cm path. It is reasonable to take this fraction as the fractional error in radius of curvature produced by coulomb scattering. It is a number which increases with decreasing particle velocity and magnetic field. For the field intensity used, the error in radius due to coulomb scattering varied from 5.5 percent at the low end of the observed momentum spectrum, to 2 percent at the high end.

[14] P. M. S. Blackett and R. B. Brode, Proc. Roy. Soc. **154**, 573 (1936).

[15] B. Rossi and K. Greisen, Rev. Mod. Phys. **13**, 240 (1941).

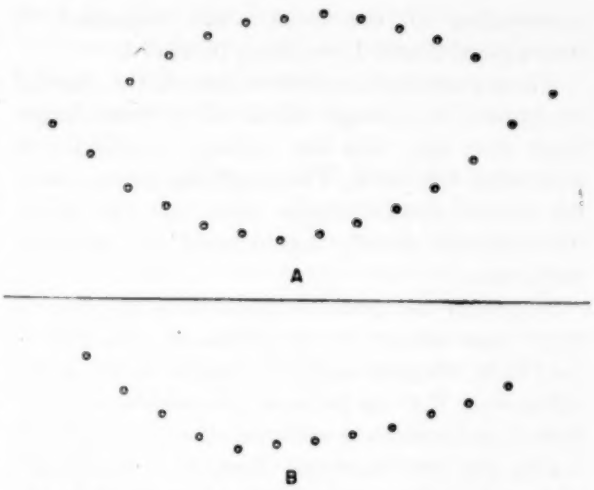


FIG. 3. In *A* and *B* are shown positive and negative tracks replotted with the vertical coordinate multiplied by ten so as to increase the apparent curvature; the tracks shown in Fig. 3(a) are Grade *A* tracks; the track shown in Fig. 3(b) is a poor Grade *B* track and shows the worst distortion allowed for a track whose apparent sign of curvature was believed.

Photography

Pictures were taken by means of a single 35 mm motion picture camera. The duration of exposure was fixed by the flash of the condenser-discharge lamps. Pictures were taken from 80 to 100 milliseconds after passage of the charged particle. Stereoscopic photographs were not taken. The probable difference between photographed and true radius of curvature due to optical effects was about 5 percent.

Analysis of Pictures

As soon as was convenient after their development, the photographs were inspected and all pictures bearing tracks which could possibly have had anything to do with the experiment were suitably marked for analysis. In this way, several percent of the pictures which bore tracks were eliminated. In general these were tracks due to small showers, knock-on electrons travelling in a markedly non-vertical direction, and tracks resulting from the nuclear radiation of neighboring laboratories.

The marked pictures were then mounted on the micrometer stage of a tool-maker's projector; they were so oriented that one of the two perpendicular micrometer screws lay parallel to the maximum chord of the circular track. The

coordinates of the track were measured at intervals of about 1 cm along its chord.

These coordinates were replotted, the chordal coordinate to a scale about 30 percent larger than true size, and the sagittal coordinate to ten times this scale. The resulting curves were, for perfect tracks, elliptic arcs (see Fig. 3(a)). This process greatly accentuated the apparent curvature.

Although the primary purpose of the experiment was simply to ascertain the qualitative fact as to whether negative mesons decay in any substance, the tracks were measured to a precision commensurate with the other experimental errors for two reasons: first, this procedure, although laborious, is alone capable of giving an objective estimate of the reliability of the apparent sign of curvature; second, it was apparent after the apparatus was in operation, that it was capable of yielding a rough measure of the mass of the particles which caused the delayed coincidences.

Two methods were tried for determining the radii of curvature of the original tracks from the plots. In the first method, the tracks were compared with calculated ellipses each of which corresponded to a circular arc of a particular radius. The track plot was bracketed between that ellipse which was just too small and the one which was just too large. The average of the two values for the radius so obtained was taken as the radius of the original track and the difference of these two radii divided by their sum was taken as indicating the *quality* of the track.

This method took too much time. It was noticed that in most of the cases in which the track departed from a perfect circle, the deficiency consisted of a continuing increase or decrease in radius along the arc (see Fig. 3(b)). The second method made use of this fact. The maximum chord was determined and the sagitta at its midpoint measured. From these, one value of the radius could be calculated. The maximum sagitta was then determined and the half-chords to the right and left of this point were measured. From these, two more values of the radius could be calculated. The average of the three values was taken as the radius of the track and the fractional average deviation of the three values from their mean was taken as a measure of the quality of the track. By measuring a given group of tracks by both methods, it was shown that they were equivalent (most of the radii were less than 80 cm).

Neither method was completely objective: the first obviously required the exercise of judgment; in the second, the small fraction of tracks which were non-circular, but still symmetric, were assigned too good a quality number and these had then to be degraded by inspection. The second method was used because it was less tedious.

After being measured, the tracks were reprojected on a true scale copy of Fig. 1(a). The direction of curvature was noted, and also the tracks were examined to make sure they were "pertinent," i.e., that they lay within the solid angle determined by the counter arrays. Non-pertinent tracks were discarded. Also discarded, were a very few slow meson pairs.

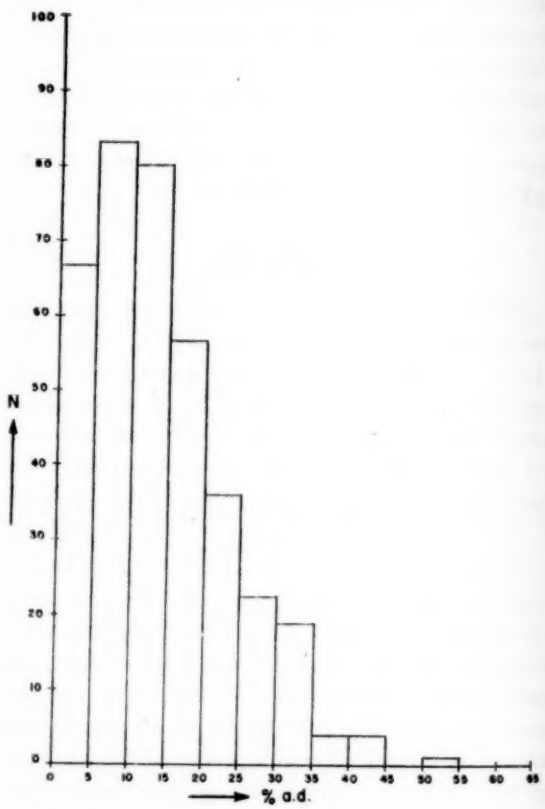


FIG. 4. Distribution of tracks in quality. The radius of curvature of each track was measured at three positions; this diagram shows the number of tracks plotted against the fractional average deviation from their mean of the three measurements on a given track. Tracks for all specimens are included.

The tracks were then graded: Grade *A* tracks are those whose quality number lies between 0 and 0.15; Grade *B* are those whose quality number is greater than 0.15 and less than or equal to 0.30; Grade *C* tracks are those whose quality number is greater than 0.30; Grade C^1 includes all those of insufficient contrast to be measured.

Tracks whose radii of curvature were greater than 250 cm were called "straight" and were not considered as being due to meson decay. This is a value considerably higher than the theoretical upper limit of the expected spectrum and it will be shown that it is also higher than the experimental upper limit. Such "straight" tracks were discarded. Figure 4 shows the distribution in quality of all the measurable curved tracks for all specimens together.

Testing Procedure

After the apparatus had operated for several days, controlled by delayed coincidences, it was stopped. The film was marked and then an additional set of pictures was taken on the same film strip, but with the cloud chamber controlled by simple four-fold coincidences between trays 1, 2, 3 and *D*. The particles whose tracks were thus photographed comprised a random sample of whatever in the cosmic radiation could penetrate the several absorbing blocks shown in Fig. 1. These particles were mostly fast mesons and thus nearly all of the tracks photographed by this second type of control had radii of curvature far greater than 250 cm.

From these simple-coincidence controlled pictures, the presence of any non-magnetic cause of track curvature, other than scattering, large enough to have invalidated the delayed-coincidence controlled pictures could easily have been detected. In addition, these pictures served to determine the recording efficiency of the cloud chamber. It will be seen from Fig. 1 that the counter tubes overhung the cloud chamber so that not every coincidence was necessarily produced by a particle which passed through the illuminated volume. The cloud-chamber efficiency was gotten by dividing the number of pictures bearing pertinent tracks by the total number of pictures taken, in a given run. A similar efficiency ratio could of course be gotten from the delayed-coincidence controlled pictures. Whenever these two values of efficiency departed widely from one another, the circuits were suspected of malfunction. Actually this happened only once.

The general idea is useful, however, because it is possible to study the performance of the apparatus as a whole when it is controlled by simple four-fold coincidences since they occur at the rate of about 500 per hour. Because of the few delayed coincidences (less than one per hour) the delay circuits can only be tested by artificial pulses. This, then, is a way to help ascertain that the cloud chamber has functioned properly over a long interval, even though at a very slow rate. At the beginning of the experiment, both efficiencies were about 45 percent. By improving the lighting system and the pressure-temperature controls, this was gradually increased to 75 percent, which is approximately the limit set by the counter geometry.

Now when the apparatus was controlled by delayed coincidences, there were recorded a certain number of "straight" tracks which were certainly not due to slow mesons. These averaged to 15 percent of the pertinent tracks recorded by delayed coincidence control. However, when the apparatus was controlled by simple four fold coincidences, more than ninety percent of the tracks were due to fast particles. This is additional proof that the apparatus functioned properly since the delayed coincidence control increased the relative number of slow particles about fifty fold. These numbers also show that only 1.5 percent of the slow particles whose tracks were photographed during delayed-coincidence control were not germane. The presence of the "straight" tracks has therefore no important bearing on the results.

Two additional tests were made. The duration of the long-gating pulse of the delay circuit was increased from 9 μsec. to 17 μsec. No increase in the rate at which slow meson tracks were obtained was noticed. Again, when the gating pulse was 9 μsec. in duration, the apparatus was exposed to sufficient gamma radiation to increase the counting rate of tray *D* by a factor of four. Again no increase was observed in the rate at which slow meson tracks were recorded. However, the presence of sufficient slow neutrons

from the cyclotron, to increase the individual counting rates by a factor of ten, rendered the apparatus inoperable because of the large background in the cloud chamber.

Possible Causes of Spurious-Delayed Coincidence Tracks

It is necessary for completeness to investigate why 15 percent of the tracks observed when the apparatus was tripped by delayed coincidences were "straight" (i.e., had momenta in excess of 3.5×10^8 Mev/c).

It was possible that a meson could decay in the specimen without having passed through the cloud chamber. In fact, since the measured-cloud-chamber efficiency $\eta\cong75$ percent, it was possible for 25 percent of the slow mesons to do this. It was then possible that during or immediately before the resultant expansion of the cloud chamber, an independent fast particle could traverse the cloud chamber. Since such a particle has nearly unit probability of tripping the three-fold coincidence circuit, and since most of the penetrating particles are fast, it is possible to set the three-fold coincidence rate N_3 equal to the rate at which such particles traverse the cloud chamber. Then if γ_1 is the rate at which such spurious tracks are observed,

$$\gamma_1=\eta N_3 t N_D(1-\eta)\cong.01N_D, \qquad (1)$$

where N_D is the observed rate of delayed coincidences and t is the resolving time of the cloud chamber. For the rates observed, γ_1 amounts to only about 1 percent of the delayed-coincidence rate recorded.

Another source of such tracks arises because it was possible for a particle to penetrate trays 1, 2, and 3, and the cloud chamber without triggering a tube of tray D about 10 percent of the time. Such events would have no effect on the apparatus, unless within the 9 μsec. period defined by the gating pulse of the delay-coincidence circuit, a tube of tray D was triggered by another particle. The rate at which straight tracks would be observed due to this cause is,

$$\gamma_2=\eta(N_3-N_4)n_D\tau, \qquad (2)$$

where N_4 is the four fold coincidence rate, n_D is the separate counting rate of tray D and τ is 9 μsec. This can amount to less than 2 percent of delay coincidences recorded if there is no non-cosmic radiation. A recording counting-rate meter was kept in operation continuously to check the background of non-cosmic radiations. This indicated that this effect could actually have amounted to 3.5 percent of the delay coincidences recorded.

The imperfect anticoincidence efficiency can in principle contribute in another way to the spurious tracks recorded, for a moderately fast meson can penetrate the specimen without triggering the counters of tray D about 10 percent of the time. It can then decay in the material of the supporting framework and send back an electron to trigger tray D, thus establishing a true delayed coincidence. Only mesons whose range after penetrating the specimen S is less than that of a decay electron can do this. Solid angle calculations indicate the effect to be small. In this experiment the supporting framework was of wood. Since few if any negative particles appeared to decay in brass or in iron this estimate is confirmed because the wood would have allowed mesons of either sign to decay.

The charging-time constant for a subgroup of tubes which shared a common charging resistor and coupling condenser was at most 50 μsec. Now if one tube of such a subgroup was fired, the voltages applied to the other tubes of that subgroup were so lowered that they would yield smaller than normal pulses until the 50 μsec. circuit was recharged. The preamplifier-output pulses which ensued then rose to their limited maximum value more slowly. It is not probable that any of these abnormally low pulses contributed a delay of as much as one microsecond. However, even if it is assumed that the output pulses were delayed by one microsecond for a period of 100 μsec. after discharge of one tube, only about 0.3 percent of the delayed coincidences observed could be due to this cause.

After a self-quenching counter tube has fired, it is known to require several hundred microseconds to recover sufficiently to fire normally a second time. During approximately the last 100 μsec. of this "dead-time," the tube will give out pulses, but of less than the normal size.[16] If it is assumed that all of these abnormally small

[16] A. Trost, Zeits. f. Physik **105**, 399 (1937).

pulses contribute fictitious delays of one microsecond or more, this effect could contribute at most 0.3 percent of the observed delayed coincidences.

Aside from the above effects, the pulses from counter tubes do not issue immediately after passage of the charged particle but may be delayed on the average as much as several tenths of a microsecond. Mr. M. L. Sands has measured these spontaneous delays for tubes similar to the ones used in this experiment. He finds that not more than one pulse in 5×10^3 can be delayed by as much as one microsecond. If this finding is interpreted to mean that one in five thousand of the pulses actually is delayed by one microsecond, then 10 percent of the delays recorded would be fictitious due to this cause.

It is concluded that spontaneous counter delays and the effect described by Eq. (2) are chiefly responsible for the "straight" tracks.

Theoretical Limits of the Momentum Spectra

The arrangement shown in Fig. 1 imposes both an upper and a lower limit to the momentum spectrum of the observed tracks. No meson which passes through the specimen S and one of the counters D and then decays can cause a true delayed coincidence. This is so even if the decay electron comes back and triggers a D counter which is connected to a different preamplifier pentode, since the blocking oscillator peaking circuit in the delay-coincidence set has a recovery time of about 25 μsec., during which it cannot be triggered successfully. Thus the particle whose momentum is measured in the cloud chamber has a maximum permissible range. Therefore there is a maximum observable momentum for a given mass. The maximum amount of material which can be traversed includes the lower wall of the cloud chamber, water jacket J_2, the walls and shield can of tray 3, the specimen and, for positive mesons, the upper wall of the long counters of tray D. Obviously, this is to be computed for the trajectory which makes the greatest permissible angle with the vertical. There are in addition, a few particles which travel almost entirely in the brass walls of the vertical portions of tray D. These have a much longer permissible maximum range. For this reason and because of scattering within the specimen, the upper limit is not expected to be sharp. The *lower* limit to the spectrum is set by the fact that it is necessary to discharge a counter of tray 3 in coincidence with trays 1 and 2 before a delayed pulse from tray D can have any effect. Therefore, any meson whose momentum is measured in the cloud chamber must have at least enough energy to penetrate to the active volume of one of the counters of tray 3. Such a meson, if positive, must be able to penetrate the wall of the cloud chamber, the water jacket J_2, the upper part of the shield can of tray 3 and one wall of a tube of tray 3. If the meson is negative, it must penetrate the lower walls of tray 3 in addition, because it cannot disintegrate in brass. The lower limit to the momentum spectrum also depends on the meson mass and is not expected to be broadened by the factors which broaden the upper limit. This lower limit is not dependent upon the specimen used.

Now one would expect the momentum spectrum to fall at the low end anyhow if the original spectrum is uniform in range, because of the parabolic nature of the mean range-momentum law in the non-relativistic region.[17,18] It will be shown that the observed cut-off is much sharper than this, however. Furthermore, the minimum radius of a particle which can traverse the magnetic field and still trip all the counters is at least three times smaller than the minimum radius observed, so that the observed lower limit is not due to this cause.

TABLE I. Tracks of each category observed when specimen was carbon.

Sign	Grade *A*	Grade *B*	Grade *C*	Grade *C*'	Non-perinent
+	49	30	7	5	6
−	34	22	13	8	3

III. RESULTS

Except for a daily test period, the apparatus was operated continuously for periods of several weeks duration with a given specimen material.

In the first experiment, the specimen was composed of carbon. This experiment indicated

[17] G. C. Wick, Nuovo Cimento **XXI**, 1 (1943).
[18] J. H. Smith, Phys. Rev. **71**, 32 (1947).

that negative mesons decay, in accord with the findings of Conversi *et al.* The apparatus was then refined, particular attention being given to improving the illumination and temperature control of the cloud chamber. A specimen of stainless steel was next employed and after negative mesons were found not to decay in it, the carbon experiment was repeated. It agreed with the first carbon experiment. Following this, water, brass and beryllium were tried. During part of each experiment the magnetic field was reversed. There was no possibility of the magnetic field being inadvertently reversed.

Carbon

The carbon used was in the form of pressed porous bricks whose specific gravity was 1.12. The different categories of tracks of each sign obtained when the specimen was composed of carbon are shown in Table I, which includes the results of both carbon experiments.

The ratio of positive to negative mesons of Grades *A* and *B* together was, when the magnetic field was directed toward the camera 33/14, and when the field was directed away from the camera 46/42. The dependence of this ratio upon the direction of the magnetic field was due to the fact that counter tray 3 was displaced horizontally (Fig. 1(a)). This was rectified before the water experiment was begun, and then no such dependence was discernable. The desired ratio is easily seen to be the geometric mean of these two values. It must be further corrected because the incident mesons are known to

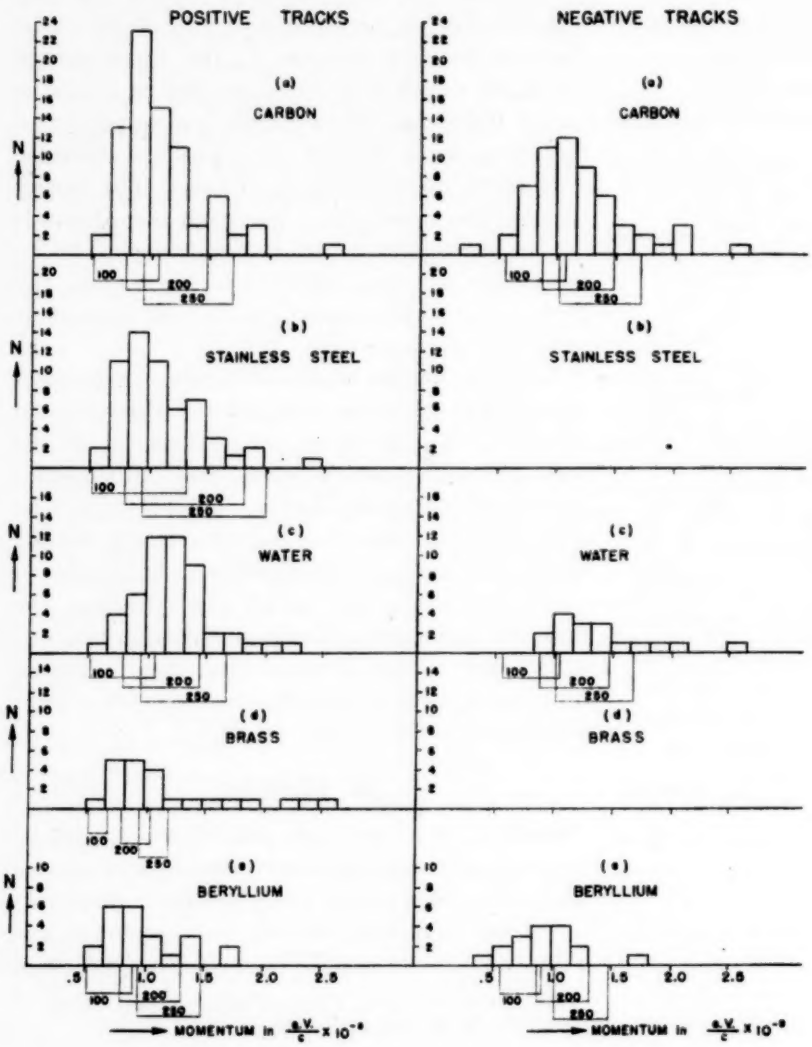


FIG. 5. Momentum distributions of positive and negative tracks for various specimens. The rectangular brackets underneath each distribution show the location and width of the main part of the expected distribution if the meson mass is 100, 200 or 250 times that of an electron. The ordinates are not corrected for the effect of the counter walls or of the copper can.

TABLE II. Tracks of each category observed when specimen was stainless steel.

Sign	Grade *A*	Grade *B*	Grade *C*	Grade *C*¹	Non-pertinent
+	42	17	10	12	1
−	0	0	1	1	1

comprise about 20 percent more positives than negatives.[19,20]

The observed positive to negative ratio, thus corrected, is for carbon: 1.33 ± 0.19. This does not give the relative number of positive and negative mesons which decay in carbon however, because some of the positive mesons could decay in the counter walls. The mass per unit length in the horizontal direction (Fig. 1(a)) was for carbon 82 gms/cm. and for the counters 49 gms/cm.; then if k is the fraction of negative mesons which disintegrate, we should set

$$(82+49)/82k=1.33\pm0.19$$

on the assumption that the efficiency for detecting decay electrons from the carbon and from the brass is the same. Then $k=1.2\pm0.17$. Since the maximum expected value for k is unity, the experiment indicates that all negative mesons decay in carbon and that none undergo nuclear capture.

Figure 5(a) shows the momentum distribution of the Grade *A* and Grade *B* tracks together with brackets which indicate the expected spectrum limits for several values of the meson mass. The registered delay-coincidence rate was 0.74 ± 0.04 per hour for 493 hours of operation. There were photographed 173 tracks of Grades *A*, *B*, and *C* together and also 22 "straight" tracks. The register data should be corrected for background, therefore, by 13 ± 3 percent. This gives for the counting rate 0.65 ± 0.04 per hour.

Figure 6 shows two contiguous pictures on the same film strip. It will be noted that the tracks curve in opposite directions.

Iron

The iron specimen comprised ten pieces of non-magnetic stainless steel each $\frac{1}{8}$-inch thick and separated from one another by $\frac{3}{8}$ inch. The specimen was of about the same external dimensions as the other specimens (carbon, water, beryllium) but had somewhat more than twice their average density. Non-magnetic steel was used in order not to affect the magnetic field. Its composition was 73.8 percent iron, 18 percent chromium, 8 percent nickel, and 0.2 percent carbon. The mass of the specimem per unit length was 210 gms/cm.

The different categories of tracks of each sign

FIG. 6. A positive and a negative meson track photographed on adjacent frames of the same film strip (carbon specimen). The sharply curved negative track at the top of the upper picture is known not to be contemporaneous with the meson track since its narrow width indicates that that particle passed through the cloud chamber after the expansion.

TABLE III. Tracks of each category observed when specimen was water.

Sign	Grade *A*	Grade *B*	Grade *C*	Grade *C*¹	Non-pertinent
+	37	17	1	2	2
−	10	8	3	2	1

[19] H. Jones, Rev. Mod. Phys. 11, 235 (1939).
[20] D. J. Hughes, Phys. Rev. 57, 592 (1940).

obtained when the specimen was composed of stainless steel are shown in Table II.

It is obvious that very few, if any, negative mesons decay in stainless steel. The two Grade *C* tracks can easily be accounted for as having accompanied the 22 "straight" tracks which were also observed.

The registered delay-coincidence rate was, for 302 hours of operation, 0.74 ± 0.05/hr. The total number of Grade *A*, *B*, and *C* tracks photographed was 81, and there were in addition 22 straight tracks. From these numbers, the background is estimated to be 27 percent. Corrected for this background, the counting rate is 0.54 ± 0.06/hr.

Figure 5(b) shows the momentum distribution of the Grade *A* and Grade *B* tracks together with brackets indicating the expected spectrum limits for several values of the meson mass.

Water

Water was used in order to study the behavior of mesons stopped in oxygen. It has been pointed out† that water should behave precisely like oxygen, the only effect of the hydrogen being to form "hydrogen meside," an electrically neutral system which can drift within range of the oxygen nuclear force in a time very small compared to the meson half-life. The water was contained in a copper can of approximately the same dimensions as the previous specimens. This can increased the mass per unit length of copper and brass to 71 gms/cm.; the mass per unit length of the water was 98 gms/cm.

The different categories of tracks of each sign observed when the specimen was composed of water are shown in Table III.

The ratio of positive to negative mesons of Grades *A* and *B* together was, when the magnetic field was directed toward the camera, 22/9; and when it was directed away from the camera, 32/9. The disagreement between these ratios is not statistically significant. The mean observed positive to negative ratio, corrected for the initial positive excess, is 2.46 ± 0.67.

Proceeding as in the case of carbon, the fraction of negative mesons which decay is found from

$$(71+98)/98k = 2.46\pm0.67$$

which gives $k=0.7\pm0.19$.

Figure 5(c) shows the momentum distribution of the tracks observed.

The registered delay-coincident rate was, for 278 hours of operation, 0.87 ± 0.06/hr. There were photographed three "straight" tracks and eighty Grade *A*, *B*, and *C* tracks. If these numbers indicate the background rate, then the corrected counting rate is 0.84 ± 0.06/hr.

Brass

This specimen consisted only of the counter tubes themselves and the empty water can. The mass of the specimen per unit length was 71 gms/cm.

The different categories of tracks of each sign observed for this specimen are shown in Table IV.

The registered delay-coincidence counting rate was, for 160 hours of operation, 0.26 ± 0.04 per hour. A total of 23 tracks of Grades *A*, *B*, and *C* were photographed as well as two "straight" tracks. If these numbers indicate the background rate, the corrected delay coincidence rate was 0.24 ± 0.04/hr.

Figure 5(d) shows the momentum distribution of these tracks. Note that it has the narrow peak and long high momentum tail which would be expected from a hollow specimen.

Beryllium

The can previously used to hold water was packed with beryllium chips. The mass per unit

TABLE IV. Tracks of each category observed when specimen was brass.

Sign	Grade *A*	Grade *B*	Grade *C*	Grade *C*[1]	Non-pertinent
+	19	2	2	0	1
−	0	0	0	0	0

TABLE V. Tracks of each category observed when specimen was beryllium.

Sign	Grade *A*	Grade *B*	Grade *C*	Grade *C*[1]	Non-pertinent
+	9	14	7	4	0
−	11	6	1	0	0

† At the Conference on the Foundations of Quantum Mechanics held at Shelter Island, New York, June 1–3, 1947.

TABLE VI. Comparison of delayed coincidence rates.

Specimen	Carbon	Stainless steel	Brass and copper	Water	Beryllium
Rate per hr per gm/cm	0.0059±0.0003	0.0018±0.0002	0.0034±0.0005	0.0061±0.0007	0.002±0.0015

length of beryllium was 70 gms/cm. and of brass and copper 71 gms/cm.

The different categories of tracks and each sign observed in this experiment are shown in Table V.

The positive to negative ratio corrected as in the cases of carbon and water for the initial positive excess is 1.13±0.36. The fraction of positive mesons which decay is found from

$$(70+71)/71k = 1.13 \pm 0.36,$$

which gives $k = 1.76 \pm 0.56$. Since the maximum possible value of k, if all positive and all negative mesons decay, is unity, this figure does not indicate that any negative mesons are captured.

Figure 5(e) shows the momentum distribution of the beryllium tracks.

The registered delayed-coincidence rate was, for 369 hours of operation, 0.61±0.04/hr. The total number of Grade *A*, *B* and *C* tracks photographed was 48, and in addition 18 "straight" tracks were recorded. If these numbers indicate the background, the corrected rate was 0.38±0.26/hr.

Comparison of Counting Rates

By subtracting the counting rate for brass from those for the water and beryllium experiments, we arrive at the counting rate for the water and beryllium alone. Since the copper can was not used during the carbon and iron experiments, only 49/71 of the brass counting rate needs be subtracted from their rates to get the rates for carbon and iron alone. If each of these rates is divided by the linear density of its respective specimen, the resultant ratios should be comparable in the sense of Sigurgiersson and Yamakawa.[13] The results are shown in Table VI.

It should be remarked that the apparatus behaved poorly during the beryllium experiment which was terminated by its breaking down. The two sets of data for beryllium, if accepted at their face value, would indicate that positive mesons undergo nuclear capture in this element. Because of the known deficiencies of the apparatus during the time beryllium was being tested, this is not regarded as conclusive and the matter is being subjected to further test.

The results for carbon, steel, brass and water tend to validate the method used by Sigurgiersson and Yamakawa, especially in view of the great difference in electron absorption to be expected for the steel and for the brass specimens, due to their different geometries.

Professor Bruno Rossi suggested this problem. The writer is indebted to him for his constant advice and encouragement.

Mr. M. L. Sands was associated with the work during its early stages.

Much of the apparatus was constructed by Mr. R. Hewitt, technician. Mrs. T. Kallmes measured all the tracks and prepared plots of which Fig. 3 is an example.

The research reported in this document was made possible through support extended the Massachusetts Institute of Technology, Laboratory of Nuclear Science and Engineering, by the Navy Department (Office of Naval Research) under Contract N5ori-78, Task Order VI.

Discharge Spread in Geiger Counters

I. With Self-Quenching Gases

J. D. CRAGGS AND A. A. JAFFE
High Voltage Laboratory, Research Department, Metropolitan-Vickers Electrical Company, Ltd., Manchester, England

(Received June 13, 1947)

Despite the work of Korff, Stever, and others, many aspects of Geiger counter behavior are as yet unexplained. In particular, the extent to which the photons produced in the counter discharge are of importance, and the manner in which they are absorbed in the counter gas, seem to have been insufficiently studied. The experiments described here were made with single wire, divided cathode, counters of various types. The results show that photon absorption seems to be less complete than is often thought, but that when such absorption is fairly small, it may be almost entirely manifest as a cathode, not a gas, effect, as shown by experiments with a new type of divided cathode counter in which the cathode effect is virtually eliminated.

I. INTRODUCTION

THE differences in the behavior of proportional counters containing simple and complex gases[1,2] have been attributed mostly to the greater importance of photo-emission of electrons from the counter cathode in the former case. The experiments of Rose and Korff[2] support this hypothesis. In the case of non-proportional (i.e., Geiger-Müller) counters the experimental work does not seem to have shown either the absence or presence of the cathode photo-emission with any certainty.

Stever,[3] Wilkening and Kanne,[4] Nawijn,[5] and others have investigated the effects of various barriers, e.g., glass beads sealed on to the anode wire, on the spread of the discharge along the counter tube, following earlier similar work by Greiner[6] on hydrogen and by Curran and Strothers[7] and Swiss workers[8] on various gas mixtures. It is easy to conclude from these and other experiments that the counter discharge spreads much less in the self-quenching type of counter, containing a complex (polyatomic) gas, than in the non-self-quenching (simple gas) type containing, for example, argon or hydrogen. It has even been stated[9] (see the Discussion below) that there was no evidence, in self-quenching counters, for the presence of photons at appreciable distances from the wire, although most authors do not take this extreme view. It was considered worth while to investigate the discharge spread by photo-ionization in counters more thoroughly, with particular emphasis on such simple gases as argon and hydrogen. The first part of the work was on self-quenching mixtures, and the results are reported here since it is hoped they may be of interest to users of counters containing such mixtures. It was thought necessary, for completeness, to mention the work with non-self-quenching gases in the present paper.

II. APPARATUS

Divided cathode, single anode, counters of the type used for other experiments by Stever,[3] Ramsey and Hudspeth,[10] and others were employed for some of the experiments. They consisted of either 6 (Fig. 1) or 12 cathode assemblies, the copper cathodes being 2 inches or 1 inch long, respectively, 0.75 inch internal diameter, spaced $\frac{1}{8}$ inch apart. The anode wires were of 5-mil tungsten. The cathodes were given various chemical treatments (see below); the pumping

[1] S. A. Korff and R. D. Present, Phys. Rev. **65**, 274 (1944).
[2] M. E. Rose and S. A. Korff, Phys. Rev. **59**, 850 (1941).
[3] H. G. Stever, Phys. Rev. **61**, 38 (1942).
[4] M. H. Wilkening and W. R. Kanne, Phys. Rev. **62**, 534 (1942).
[5] A. Nawijn, *Het Gasontladings Mechanisme Van Den Geiger-Müller Teller* (Drukkerij Hoogland-Delft, 1943), pp. 52, 53.
[6] E. Greiner, Zeits. f. Physik **81**, 543 (1933).
[7] S. G. Curran and J. E. Strothers, Proc. Camb. Phil. Soc. **35**, 654 (1939).
[8] F. Alder, E. Baldinger, P. Huber, and F. Metzger, Helv. Phys. Acta **20**, 73 (1947), and references therein to earlier work.
[9] P. Weisz, Phys. Rev. **62**, 477 (1942).
[10] W. E. Ramsey and E. C. Hudspeth, Phys. Rev. **61**, 95 (1942).

technique was not good, but was of the simple nature described and discussed, for example, by Rochester and Janossy[11] and found to be quite satisfactory for the production of reliable self-quenching counters.

Two other types of counter were used. In the first of these (Fig. 2) which contained 4 copper cathodes (above dimensions), the arrangement was as shown. The second new type of counter had 6 cathodes (above dimensions) spaced 0.5 inch apart as in the counter of Fig. 2, with a window (as in Fig. 2) attached to the front of each cathode except the first; it is shown schematically in Fig. 3. Experiments with such window-cathode counters, designed to eliminate the cathode effect by collimating the photon beams to the neighborhood of the wire, do not appear to have been previously described. The device, due to Wilkening and Kanne,[4] of inserting a Bakelite ring in the cathode is only superficially similar. Collimation is not perfect in cylinders 1 and 2 of Fig. 3, since the discharge region in 2 is visible at the cathode of 1. Great care was taken in the lining-up of all the cylinders in the multi-cathode counters.

The method of experiment was to irradiate an extreme cathode by carefully collimated gamma-rays from radium, and to determine the extent of discharge spread by counting coincidences between the irradiated cathode and each of the other cathodes in turn. The voltage on intervening cathodes was reduced to a value just below the starting potential, to avoid effects attributable to spurious charges and to prevent diffusion of electrons along the counter. Many check experiments to assess the importance of these and other effects were made.

The coincidence circuit is shown in Fig. 4. The mixer valves are preceded by SP.41's. The scaler and counting circuit will be described elsewhere. The amplification given by the pulse

FIG. 1. 6-cylinder counter.

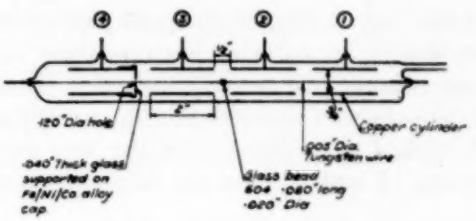


FIG. 2. 4-cylinder counter.

reversing stages of Fig. 4 was valuable since the pulse size, especially from the 1-inch cathodes, was fairly small. When coincidences with non-self-quenching gases were to be counted, the quenching circuit of Fig. 5 was used; it is similar to that of Neher and Harper.[12] It was, of course, also necessary in our case to arrange for independent control of the voltage on each cathode. Careful tests were made to ensure that the coincidence circuits were symmetrical and that both channels were identical in sensitivity.

III. PRELIMINARY TESTS

In order to calculate the counting rate resulting from chance coincidences (cosmic rays, etc.) or those arising from the small amount of radiation passing sideways through the lead collimator block, it was necessary to find the resolving time of the circuit of Fig. 4. The method used by many previous authors was adopted here. Two small separate counters (2½-inch cathodes) were set up at a distance apart (4 inches) roughly equal to the length of the divided cathode counters. By taking individual counts (N_1N_2), using radium gamma-rays, and the chance coincidence counts in the same conditions of irradiation, etc., the resolving time (τ) may be found from the expression

$$C=2N_1N_2\tau+K.$$

N_1 and N_2 included the counter background in each case. With the circuit as used for self-quenched counters (Fig. 4), the resolving time with 0.001 μf condensers was about 30 μsec., and with 0.002 μf condensers, about 40 μsec. This could easily have been reduced but was adequate for the present work (see below). The true cosmic-ray coincidence count K may be eliminated by using two sets of data.

[11] G. D. Rochester and L. Janossy, Phys. Rev. **63**, 52 (1943).

[12] H. V. Neher and W. W. Harper, Phys. Rev. **49**, 940 (1936).

It was easy to show that circuit pick-up effects were absent, by reducing the operating voltage of one counter just below its starting potential. This was checked also with two separate counters.

A typical experiment with the two separate counters (A and B) gave the following results:

(a) with no gamma-ray irradiation:

$N_1=52$/min., $N_2=81$/min., $C=0.21$/min.

(b) with gamma-ray irradiation:

$N_1=1238$/min., $N_2=676$/min., $C=0.94$/min.

From this particular experiment τ can be calculated and found to be about 26 /μsec. The effect of the stray irradiation of the cathodes, other than that deliberately irradiated, i.e., the importance of imperfect collimation, will be assessed below, using the values of resolving time deduced above. It would have been undesirable to reduce the resolving time to a value comparable with or less than the time for discharge spread, which would be very small (10^{-8}–10^{-9} sec.) for photon-induced spread. Hill and Dunworth[13] have measured the velocity for discharge spread by electron movement, etc., and find it to be $\sim$10 cm/μsec.

IV. EXPERIMENTAL RESULTS WITH DIVIDED CATHODE COUNTERS

Typical results with a 6-cathode counter (Fig. 1) are shown in Fig. 6 for various self-quenching gases. Even for a case of little discharge spread (1 cm Hg ethyl ether/10 cm Hg argon) there were still 16 discharges spreading per 5 min. over 2 inches of counter with 3000 discharges per 5 min. in the irradiated counter. The curves of Figs. 6 and 7 should be corrected (before absorption coefficients are deduced) for the lack of perfect collimation, but the lack of knowledge of the effective size of the cylinder of radiation in the initiating cylinder makes this a questionable procedure. The results are therefore meant to be comparative only; further data in collimated tubes will be obtained.

The counter conditions for these curves were as follows:—

The operating voltage was generally about 40 volts above the starting voltage. The extent of the stray irradiation from the collimator may be judged by the following representative results, taken with 1 cm Hg ethyl ether/10 cm argon. The counts on the first and third cylinders (2.25 inches apart, i.e., one cathode length+two gap lengths) individually were 499 and 35/minute, and the coincidence rate was 13/minute. The chance coincidence rate, calculated as above from a knowledge of resolving time was 0.02/minute, with a true coincidence count caused by cosmic rays of about 0.2/min. The spurious coincidence rate is therefore negligible even for the most highly quenching mixtures except at the lowest counts ($\sim$1/min.). That this is a conservative estimate is shown by the repeated coincidence counts of about 0.5/min. that were obtained for the two extreme cathodes in the tube, despite the presence of the stray irradiation from the second to the last cathode (typical variations in the latter being from 35/min. to 25/min. for the second and last cathodes, respectively, with the first cathode counting at $\sim$500/minute). The discharge invariably spread from the first to the second cylinder (gap $\frac{1}{8}$ inch), see Fig. 1, since field penetration would be appreciable; coincidence readings were therefore only useful from the third cylinder outwards.

Check results were taken also with the 12-cylinder counter, with 1-inch cathodes in order

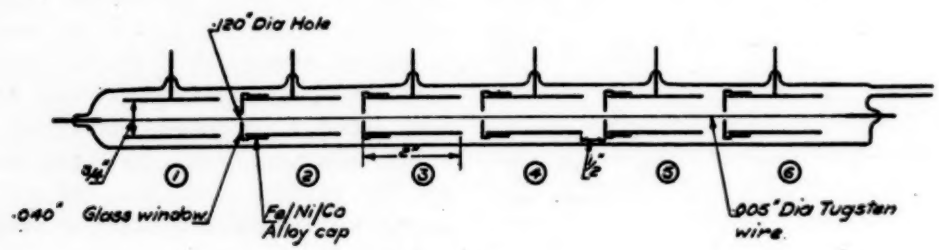


FIG. 3. 6-cylinder window counter.

[13] J. M. Hill and J. V. Dunworth, Nature **158**, 833 (1946).

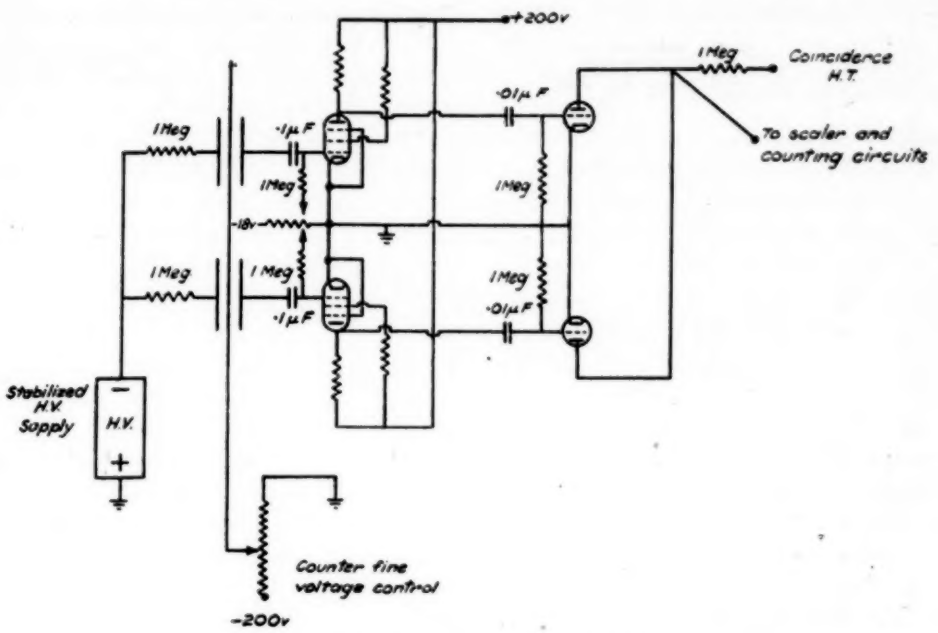


FIG. 4. 2-channel reversing stage and coincidence stage.

to obtain more accurate data for the highly quenched mixtures. Results are shown in Fig. 7 (counter operating data as in Table I).

Experiments made with hydrogen and argon showed no detectable quenching for the full counter length. The results were in accordance with expectations and are only mentioned here for completeness.

Having thus demonstrated the existence of appreciable discharge spread in counters which worked most satisfactorily in self-quenching circuits and showed excellent plateaus (flat to within ±5 percent for 200 volts) it seemed necessary to show the relative importance of cathode (photo-emission) and gas (photo-ionization) effects. Nawijn[14] considered that the discharge spread found in a beaded counter only at high voltages was due entirely to discharge spread over the bead and not to the cathode effect, since at sufficiently low operating voltages no spreading was observed. The supposition was that discharge spread caused by cathode photo-emission could theoretically have taken place at low voltages and its absence at such voltages excluded the possibility of its occurrence at higher voltages. This reasoning seems specious; in our opinion (see Fig. 9) discharge spread attributable to cathode effect is merely less likely at lower than at higher voltages. Before describing special counters made for this purpose it is of interest to quote further experimental results.

The curves of Fig. 8 show the effect on discharge spreading with two types of cathode treatment (i) heavy oxidation obtained by heating in air, and (ii) chemical cleaning with acid and washing with distilled water immediately prior to the pumping and filling. It was found that the clean cathodes gave more spreading (see Discussion, below). (Note (Fig. 8) s.v. ≡starting voltage and w.v.≡working voltage.)

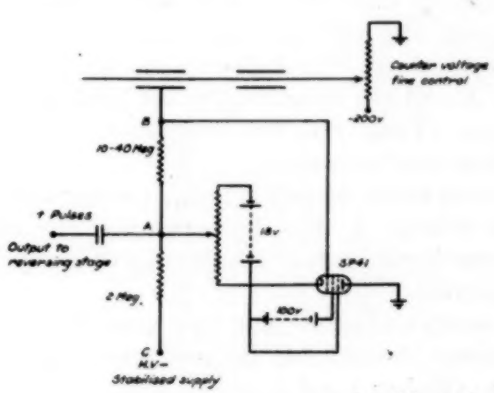


FIG. 5. External quenching circuit.

[14] See reference 5, p. 51.

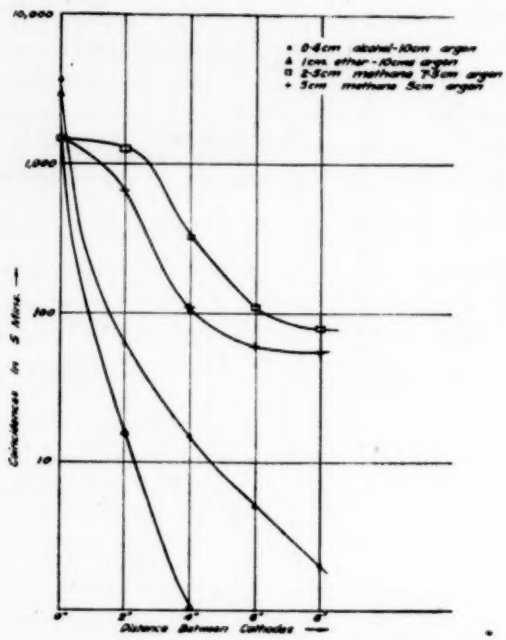


FIG. 6. Spread of discharge in 6-cylinder counter.

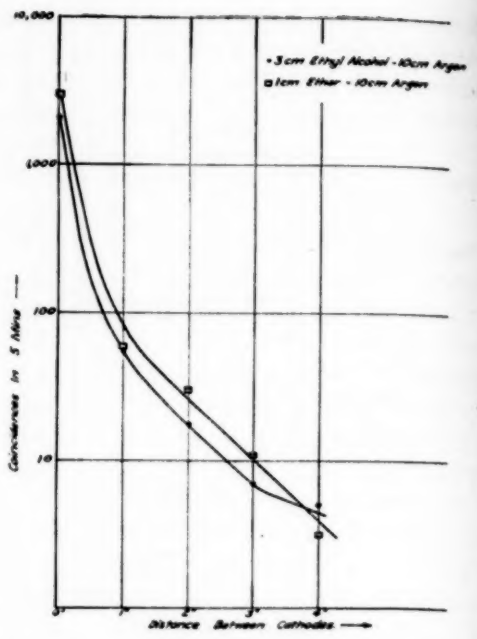


FIG. 7. Spread of discharge in 12-cylinder counter.

The variation of discharge spread (between the first and third cylinders) with counter voltage and pulse size is shown in Figs. 9, 10, and 11. (1 cm Hg ethyl alcohol vapor/10 cm Hg argon.) These are representative curves; the shape of the pulse size/voltage curve has been discussed by Nawijn,[14] and the first part only of it by Korff and Present.[1] Further and more detailed experiments on the significance of these results will be performed.

V. EXPERIMENTAL RESULTS WITH SPECIAL COUNTERS

Some experiments with the 4-cathode counter of Fig. 2 were made with the intention of assessing the relative importance of gas and cathode effects. The cylinders are numbered in Fig. 2, and the mean results are tabulated for brevity (Table II). The irradiated cathode is marked with an asterisk.

It was found, by taking plateau characteristics with cylinder 4, that the presence of the glass window seemed to have no effect on the counter's properties.

Tests with argon and hydrogen (5 cm Hg pressure) showed 100 percent spreading even with cylinders 3 and 4.

The results of Table II may be summarized briefly by stating that the spread of discharge for the simple divided cathode counters is very largely due to emission at the cathodes due to photon bombardment. The discharge spread with the window counter is much less than with the beaded counter (cylinders 2 and 3 of Fig. 2). If the discharge spread with the beaded counter was due entirely to cathode effect, the latter being unaffected by the presence of a bead, then the sum of counts for cylinders 2 and 3, and 3 and 4 (Fig. 2) should equal the number for 1 and 2. That this is not so, in general, is probably due to the above assumptions failing; that they do so to varying degrees in different gas mixtures (Table II) is interesting and, we intend, will be studied further. The effect of a possible amount of electron diffusion between cylinders (Fig. 2) was not ignored; this effect was, we consider,

TABLE I. Counter conditions for measurements with divided cathode counters.

Gas mixture	Starting voltage approx.
0.4 cm Hg alcohol/10 cm Hg argon	750
1 cm Hg ether/10 cm Hg argon	1130
2.5 cm Hg methane/7.5 cm Hg argon	1370
5 cm Hg methane/5 cm Hg argon	1500
3 cm Hg alcohol/10 cm Hg argon	1320

satisfactorily eliminated in the other divided cylinder counters (Figs. 1 and 3).

To check the effect of windows further and to obtain more accurate results on discharge spread through photo-ionization, the counter with 5 windows and 1 normal cathode was used. The alignment of the 5 windows was carried out with great care, and checked before the wire was inserted and the tube sealed at the ends. Results with a mixture of 0.5 cm Hg ethyl alcohol/10 cm Hg argon (starting voltage 760), which was known to give relatively poor quenching, gave at 800 volts 15 coincidences in 5 min. between cylinders 1 and 2 counting at 751 and 44/min., respectively. Between cylinders 1 and 3 (756 and 49/min.) there were no coincidences in 5 min., while there would have been about 50–60 coincidences in that time with the cathode effect operative (Figs. 6 and 7). It therefore seemed safe to conclude that the discharge spread in the latter figures was due almost entirely to the cathode effect in self-quenching gases. As a check, some tests were made with a mixture of 2.5 cm Hg methane/7.5 cm Hg argon, since this had given the largest amount of discharge spread in the previous experiments (Fig. 6). The starting voltage was 1300, and at 1340 volts there were 1402 coincidences in 5 minutes between cylinders 1 and 2, counting at 798 and 317/min., respectively. Between cylinders 1 and 3 there were 3 coincidences in 5 minutes, showing again that gas quenching is most efficient, with cylinders 1 and 3 counting at 789 and 42/min. Absorption curves for methane (10 cm Hg) and with ethyl alcohol/argon (0.4 cm Hg/10 cm Hg) taken with the counter of Fig. 3 are shown in Fig. 12, and these curves should represent the absorption of photo-ionizing radiations from the irradiated end cylinder. It is possible to deduce the true absorption coefficient for the operative photons from Fig. 12 by allowing (inverse square law) for the lack of perfect collimation. Thus the curve shown in Fig. 12 (crosses) is derived, and it is seen that the observed absorption effect cannot be due to the inverse square fall-off of radiation. It should be emphasized that the absorption even in methane is so large that the coefficient cannot accurately be derived from Fig. 12. Further experiments will be performed with another counter tube incorporating a collimation chamber.

VI. DISCUSSION OF RESULTS

It has, apparently, generally been supposed that discharge spreading in self-quenching counters by virtue of photo-ionization in the gas, or by liberation of electrons from the counter cathode due to the incidence on it of photons, is negligible or indeed absent. For example, Collie and Roaf[15] used a double counter built for β-ray spectrographic work and reported that the number of coincidences (12/min.) barely exceeded the number (10/min.) obtained with a separate pair of counters. They conclude, since the passage of electrons between counters might account in their case for this difference, that the photo-effect at the cathode is much too small to have any influence on discharge maintenance.

Stever[3] carried out exhaustive tests on divided counters of various types but did not use a window counter (such as that shown in Fig. 3).

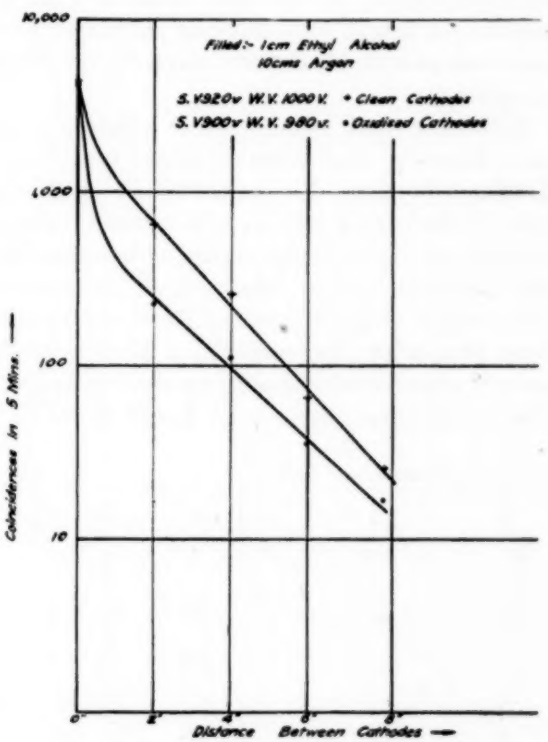


FIG. 8. Effect of cathode treatment on discharge spreading in 6-cylinder counter.

[15] C. H. Collie and D. Roaf, Proc. Phys. Soc. 52, 186 (1940).

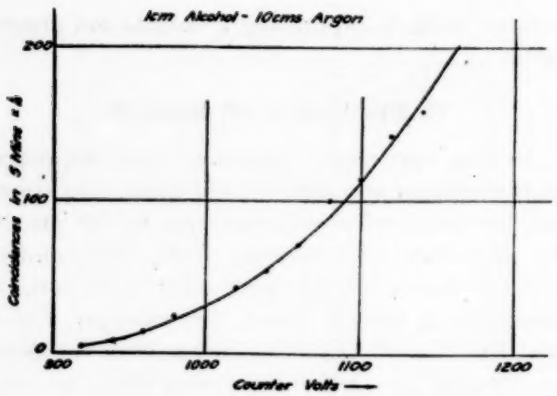


FIG. 9. Coincidence rate as function of counter voltage.

He found no spreading with a $\frac{1}{8}$-inch diameter bead, $\frac{1}{4}$ inch long, on the wire, but he apparently used the NO_2 treatment of the copper cathodes which is known to produce a surface of high work function. With a smaller bead at a lower counter pressure the discharge could be made to spread; it is not certain, however, that even then the discharge spread was due purely to photo-ionization in the gas, since a cathode mechanism might have become operative with the lower photon absorption given by the lower gas pressure.

Wilkening and Kanne[4] have published (as have Ramsey[16] and others) similar results and Korff and Present,[1] discussing their results in the light of the comprehensive and masterly theory of Korff on fast counter action, stress also that fast counters showed 100 percent localization. This is misleading, and indeed Korff and Present state that with, for example, a 50 percent/50 percent argon/methane mixture there is appreciable photo-effect. Rose and Korff[2] have, in a well-known paper, shown the presence of a cathode effect in certain fast gas mixtures used in proportional counters. However, such a counter would work in a self-quenching arrangement, and it therefore seems that the existence of an appreciable cathode effect is not a serious objection in a fast non-proportional counter. We believe this to be reasonable since the essential part of the fast counter mechanism is that first explained by Korff,[1] namely, the absence of secondary emission from bombardment of the cathode by the positive-ion sheath in such a counter; this seems to be quite explicable on qualitative but sound physical grounds.

Weisz[8] stated that no evidence has been found for the presence of photons at appreciable distances from the wire. This may be the case, with cathodes of very high work function, but Weisz's statement appears to us to be incorrect, since no qualification is made.

Korff[1] appears to dismiss rather summarily the presence of photons incapable of absorption in the complex constituent in, say, an alcohol/argon counter. The present experiments certainly seem to support that hypothesis since the discharge spread is almost zero (Fig. 6) for quite short, but appreciable (~10 cm), distances in suitable gas mixtures.

Objections have been made that the presence of appreciable photo-emission at the cathode is inconsistent with Simpson's[17] pulse-reversal method of shortening dead times by attracting the positive-ion sheath into the wire, since the continued emission of electrons would prevent the production of a short clearing time. It is

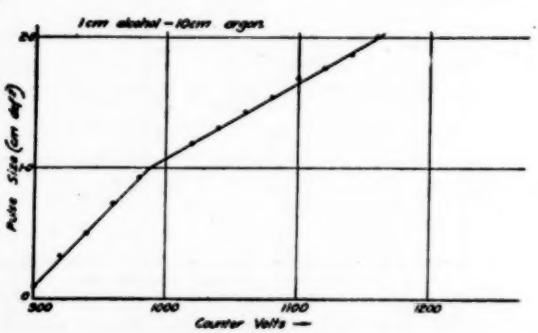


FIG. 10. Variation of pulse size with counter voltage.

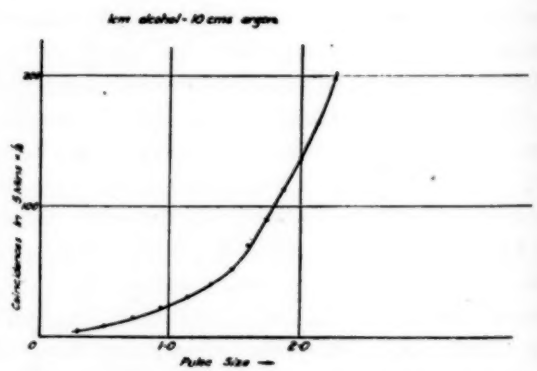


FIG. 11. Variation of coincidence rate with pulse size.

[16] W. E. Ramsey, Phys. Rev. 61, 96 (1942).

[17] J. A. Simpson, Phys. Rev. 66, 39 (1944).

TABLE II. Measurements on the importance of gas and cathode effects.

Gas filling 10 cm Hg argon with	Starting voltage	Working voltage	Counting rates with cathode (per min.) 1	2	3	4	Coincidence rate (per min.) 1	2	3
0.4 cm Hg ethyl alcohol	approx. 840	860	48	596*	—	—	81	—	—
		900	57	633	—	—	174	—	—
		860	—	52	702*	—	—	73	—
		900	—	48	707*	—	—	155	—
		860	—	—	710*	55	—	—	3
		900	—	—	707*	51	—	—	34
1.0 cm Hg ethyl alcohol	approx. 980	1000	59	803*	—	—	42	—	—
		1120	96	852*	—	—	145	—	—
		1000	—	46	899*	—	—	17	—
		1120	—	122	923*	—	—	70	—
		1000	—	—	928*	50	—	—	1
		1120	—	—	1038*	65	—	—	5
10 cm Hg methane (no argon)	2080	2140	73	716*	—	—	213	—	—
			—	36	586*	—	—	82	—
			—	—	586*	31	—	—	7

clear, however, that the existence of the reverse field entirely removes this objection and Simpson[17] indicates that photo-emission could only arise at the wire in the period during which the latter is made negative. The present experiments obviously, for that and other reasons, are not invalidated by Simpson's work.

It seemed surprising that the work function of our copper cathodes could be so high as to make them respond only to radiations absorbable in the fast counter fillings. The continuous absorption in ethyl alcohol vapor and in methane begins roughly at about 1700 and 1450A,[1] respectively, and if photons of longer wave-lengths were present in the counter discharge (so that there would be little or no absorption of them in the gas filling) then presumably the fall-off in spreading with distance (Figs. 6 and 7) would be much less than is found and would then be governed by an inverse-square law. The presence of appreciable absorption seems to suggest that the cathode work function is at least 8.5 volts

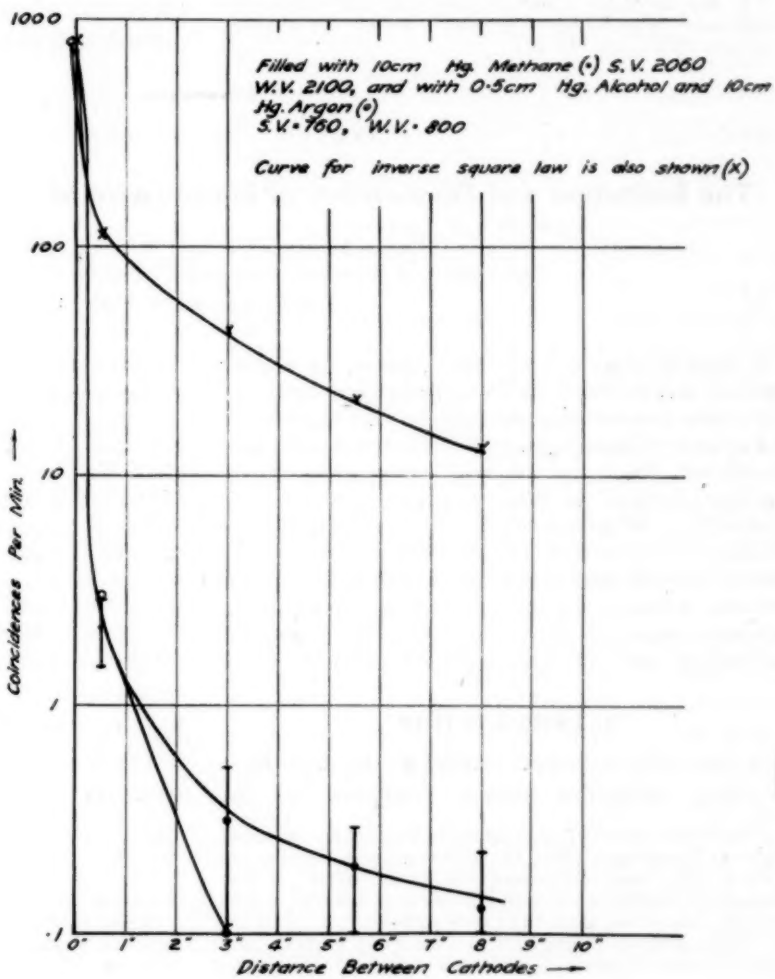

1000
100
10
1
·1
Filled with 10cm Hg. Methane (•) S.V. 2060
W.V. 2100, and with 0·5cm Hg. Alcohol and 10cm
Hg. Argon (•)
S.V.·760, W.V.·800
Curve for inverse square law is also shown (x)
Coincidences Per Min.
0" 1" 2" 3" 4" 5" 6" 7" 8" 9" 10"
Distance Between Cathodes

FIG. 12. Absorption curves for methane taken with the six-cylinder counter of Fig. 3.

(corresponding to $\lambda_{max}=1450$A). Apparently little data exist on the work function of oxidized copper; the work function of clean copper is 4.4 v[18] and, by analogy with the work on oxidation of Pt and Ag[19] which produced almost a doubling of the work function, a value for copper of 8.5 volts seems possible and reasonable.

Attempts made to show the difference between fairly clean and oxidized copper by a rough vacuum technique which was, however, as good as that often used in the preparation of satisfactory counters,[10] were made, and the results showed that copper cleaned as well as possible in our apparatus had apparently a higher work function and was therefore probably better oxidized in the physical, but not chemical sense, than the grossly oxidized cathodes we prepared. It is intended to carry out further experiments with a greatly improved vacuum technique, but this deficiency of knowledge does not seriously affect the main conclusions of this paper, i.e., that counters working very satisfactorily in normal self-quenching arrangements show considerable discharge spread which is mainly due to the cathode photo-emission effect. The results of Figs. 6, 7, and 8 will therefore not be accurately reproducible; the data of Fig. 12 should have more fundamental significance.

[18] A. L. Hughes and L. A. Du Bridge, *Photoelectric Phenomena* (McGraw-Hill Book Company, Inc., New York, 1932), p. 75.
[19] J. H. de Boer, *Electron Emission and Adsorption Phenomena* (Cambridge University Press, Teddington, England, 1935), pp. 148–151.

ACKNOWLEDGMENTS

The authors wish to thank Mr. F. R. Perry for his interest in and encouragement of the work, Sir Arthur P. M. Fleming, C.B.E., D.Eng., Director of Research and Education, and Mr. B. G. Churcher, M.Sc., M.I.E.E., Manager of Research Department, Metropolitan-Vickers Electrical Company, Ltd., for permission to publish this paper.

PHYSICAL REVIEW VOLUME 72, NUMBER 9 NOVEMBER 1, 1947

The Ionization and Dissociation of Formic Acid Monomer by Electron Impact*

THOMAS MARINER** † AND WALKER BLEAKNEY
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received July 18, 1947)

In HCOOH vapor in a mass spectrometer, low velocity electrons have produced the ionized molecule, one negative and twelve positive fragment ions. For the negative ion and six positive ions the mode of production has not been established. The other fragment ions originate in the reactions requiring the least dissociation energy for the production of the corresponding radicals, although for one of these a more energetic reaction seems preferred at high electron energies. One of the ions of doubtful origin also probably originates in a reaction involving more than the minimum degree of dissociation. The first appearance potential of four ions with established reactions, of one ion of doubtful origin, and probably of another of doubtful origin, as well as the second appearance potential of CO_2^+, all indicate that the corresponding reactions involve no more than 1.0 volt excess energy. Both HCO^+ and $COOH^+$ utilize the trivalency of O^+ in their structure. $HCOOH^+$, $COOH^+$, and HCO^+ seem to have a second appearance potential close to the first. To explain this, and to make the electron impact and spectroscopic ionization potentials of the molecule compatible, it is assumed that a *trans*-form of HCOOH exists. Estimates of its abundance relative to the known *cis*-form give for the energy difference between the two forms the probable limits 1.2–3.0 kcal/mole.

I. INTRODUCTION

IN the electron-impact studies of the more than forty molecules already reported in the literature,†† the most useful information has been obtained by application of the mass spectrometer to the determination of appearance

* The data on which this paper is based were obtained prior to November, 1941; but circumstances beyond our control have delayed preparation of the paper. A preliminary presentation of data was made at a recent A.P.S. meeting (see Phys. Rev. **71**, 139(A), 1947).
** This work is a summary of Part II of a thesis presented to the Faculty of Princeton University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
† Now at the Research Laboratories of the American Cyanamid Company, Stamford, Connecticut.
†† See. H. D. Smyth, Rev. Mod. Phys. **3**, 347 (1931) for a review of work to that date, and J. A. Hipple, J. App. Phys. **13**, 551 (1942) for more recent references.

potentials of ions, that is, the minimum energies required for the production of ions in particular modes of dissociation of molecules. Since this is also true in the study of formic acid monomer, it seems desirable to review the fundamental considerations, although they have been presented more extensively elsewhere.

An appearance potential $A(X^+)$, of an ion, X^+, may be regarded as the sum:

$$A(X^+)=I(X)+P(X)+W(X),$$

where, following the usual notation,[1,2] $I(X)$ is the ionization potential of the radical, X; $P(X)$ is the energy required to dissociate the parent molecule into neutral fragments in their lowest states in the mode of dissociation giving rise to the radical; and $W(X)$ is the total excitational and mutual kinetic energy of the products of dissociation. $W(X)$ includes, in appropriate cases, electron affinities (with a negative sign) and ionization potentials of radicals other than X receiving charges in the process being considered. It is clear from the expression for an appearance potential that analysis of electron impact data is greatly dependent upon a knowledge of the energy relationships in the numerous possible reactions. Much of this can be gained from thermochemical data and from known or estimated ionization potentials and electron affinities. When this knowledge, used in conjunction with the measured appearance potential of an ion, is adequate to establish in what reaction the ion was produced, other data may be derived. Ionization potentials of molecules themselves, which can be determined uniquely and usually within closer limits than derived data, are likely to be subject to five or ten times the error of similar quantities obtained by accurate analysis of band spectra. Derived ionization potentials of radicals will generally be subject to even greater error, and derived thermodynamic data will be quite inaccurate on chemical standards; but when the quantities have been, or can be, determined only in this way, the values are of interest.

[1] L. G. Smith, Phys. Rev. **51**, 263 (1937).
[2] C. S. Cummings II and Walker Bleakney, Phys. Rev. **58**, 787 (1940).

II. EXPERIMENTAL

Apparatus

The mass spectrometer used in the study of formic acid monomer was designed particularly to achieve high resolving power for other purposes and was only incidentally applied to the present problem. Since the instrument will be fully described elsewhere, only a brief description will be presented here for the purpose of giving significance to the conditions of operation prevailing during the experiments on formic acid monomer.

Crossed electric and magnetic fields were used to perform the m/e analysis as in the mass spectrometer described by Bleakney and Hipple,[3] who have discussed the properties of this arrangement. Differences in mechanical design, necessitated by the larger size of our instrument, are unimportant so far as the present work is concerned. The ion orbit was a prolate cycloid with $b=262$ mm, $\theta=89.5°$ (in the notation of ref. 3), and the ratio of ion-accelerating voltage to analyzing-electric field = 8.8 cm. A magnetic field of about 2230 gauss was used, except that for observation of H^+ and H_2^+ it was reduced to about 540 gauss. For the electron-impact experiments the entrance- and exit-slit widths were 0.05 mm and 3.18 mm, respectively, so that the resolving power was only 81. Such low resolution was used intentionally to obtain broad, square-topped peaks, in order to facilitate the making of relative abundance and appearance potential measurements; and it was entirely adequate. Even at the highest sample pressures used the scattering was so slight that at $M/e=46$ the ion beam contributed less than 0.05 percent of its peak intensity to an ion beam one mass unit removed.

In order to minimize errors due to differential retardation of the electrons in the ionization chamber, no voltage was impressed between the ionization chamber and the first ion slit during appearance-potential measurements. As a check on the performance of the instrument a preliminary measurement was made of the difference between the ionization potentials of neon and argon very early in the work. It was found to be

[3] W. Bleakney and J. A. Hipple, Jr., Phys. Rev. **53**, 521 (1938).

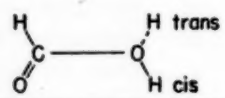
6.0±0.15 volts, in comparison with the spectroscopic value 5.78 volts.[4] In view of the fact that this determination was not made under the more favorable conditions prevailing during the experiments on HCOOH, this was taken to mean only that, at least for ions with mass-to-charge ratios near that of the calibrating gas (argon, $M/e=40$), no appreciable instrumental errors existed.

The Sample

Kahlbaum formic acid was treated in a manner similar to that found effective by Coolidge.[5] The final vapor pressure at 20°C was 32.5±0.2 mm, which, on the basis of Coolidge's data, indicates that water was not present in an amount greater than 1 percent. The sample was stored at solid CO_2 temperature except when part of it was extracted for experiments; and part of the sample was sublimed in vacuum before each extraction. The handling of the sample and the fact that Bonner and Hofstadter[6] found no perceptible decomposition of formic acid held for 36 hours at 150°C, both give assurance that no appreciable amount of impurity was present.

The sample was admitted to the mass spectrometer through a capillary leak using a gas handling system similar to that described by Smith, Lozier, Smith, and Bleakney,[7] except that no liquid air trap was used between the leak and the mass spectrometer. Pressures of formic acid between 10 and 20 mm were used on the high pressure side of the leak for appearance-potential

TABLE I. Ions produced in HCOOH by impact of 65-volt electrons.



M/e	Ion	65 volt Relative Abundance	Appearance Potential (volts)	PROBABLE REACTION	CALCULATED MINIMUM ENERGY (volts)	Deduced Ionization Potential or Additional Energy Required (volts)
46	$HCOOH^+$	100.00	11.0±0.1	HCOOH (trans?) → $HCOOH^+$	I(HCOOH) (trans?)	I(HCOOH) (trans?) ≤11.0±0.1
			11.6±0.2	" (cis?) → "	" (cis?)	" (cis?) ≤11.6±0.2
45	$COOH^+$	67.8	12.1±0.1	HCOOH (trans?) → $H-O-C\equiv O^+ + H$	I(HCOOH) (trans?) + 1.5 = 12.5 roughly	
			12.9±0.2	" (cis?) → " "	" (cis?) + 1.5 = 13.1 "	
44	CO_2^+	9.8	14.2±0.1	HCOOH → $CO_2^+ + H_2$	$I(CO_2) - .25 = 13.5$	W(X) = 0.7±0.1
			18. ±1.0	" → $CO_2^+ + 2H$	" + 4.2 = 17.9	W(X) = 0.1± 1.0
34	$H_2O_2^+$	0.06	11.6±0.4	Probably impurity	$I(H_2O_2)$	$I(H_2O_2) \leq 11.6 \pm 0.4$ probably high
33	HO_2^+	0.01	17.0±0.5	$H_2O_2 \rightarrow HO^+ + H$ from imp.	$I(H_2O_2) + 5.2 = 16.8 \pm 0.4$	
30		2.64	14.1±0.3	probably isotopic HCO^+, corresponding to higher appearance potential		
29	HCO^+	185.	12.8±0.2	HCOOH (trans?) → $HC\equiv O^+ + OH$	I(HCOOH) (trans?) + 2.3 = 13.3 roughly	
			13.2±0.3	" (cis?) → " "	" (cis?) + 2.3 = 13.9 roughly	
28	CO^+	16.7	14.6±0.3	HCOOH → $CO^+ + H_2O$	$I(CO) + 0.18 = 14.1 \pm 0.2$	W(X) = 0.5±0.5
18	H_2O^+	28.4	13.1±0.1	" → $H_2O^+ + CO$	$I(H_2O) + 0.18 = 12.77 \pm 0.05$	W(X) = 0.3± .15
17	OH^+	8.3	19.2±0.2	Indeterminate, see text		
16	O^+	3.5	21.0±0.6	" " "		
14	CH_2^+	.22	24.7±0.5	HCOOH → $CH_2^+ + 2O$	$I(CH_2) + 11.8 = 23.8 \pm 0.2$	W(X) ≥ 0.9±0.7
13	CH^+	3.0	25.4±0.5	Indeterminate, see text		
			27.4±0.6	" " "		
12	C^+	2.3	23.9±0.2	" " "		
			29.5±1.0	" " "		
2	H_2^+	.03	16.0±0.3	HCOOH → $H_2^+ + CO_2$	$I(H_2) - .25 = 15.1$	W(X) = 0.9±0.3
1	H^+	.4	19.0±0.4	Indeterminate, see text		
-17	OH^-	.03	15±5	Indeterminate		

[4] R. F. Bacher and S. Goudsmit, *Atomic Energy States* (McGraw-Hill Book Company, Inc., 1932).
[5] A. S. Coolidge, J. Am. Chem. Soc. 50, 2166 (1928).
[6] L. G. Bonner and R. Hofstadter, J. Chem. Phys. 6, 531 (1938).
[7] P. T. Smith, W. W. Lozier, L. G. Smith, and W. Bleakney, Rev. Sci. Inst. 8, 51 (1937).

measurements and slightly higher pressures for some of the relative-abundance measurements. The pressure in the ionization chamber was not determined directly; but it probably was much lower than 10^{-4} mm Hg and certainly was not above 10^{-3} mm Hg. Assuming a pressure of 10^{-3} mm Hg and a temperature not lower than 293°K in the ionization chamber, and using Coolidge's[5] equation for the dissociation constant, it is calculated that not more than 0.06 percent of the sample subjected to electron impact was in dimer form. Owing to the conservativeness of the assumptions, it is probable that the amount was even less.

Precautions Against Secondary Effects

In order to insure that the interpretation of the data would not be complicated by a high probability of secondary effects, such as having individual molecules struck by more than one electron or having individual electrons strike more than one molecule, the relative abundances of the ions were measured as a function of pressure and electron current, each variable being changed by a factor of three, encompassing the range of conditions actually used. There were found no systematic changes in relative abundances attributable to a significant probability of occurrence of the phenomena mentioned.

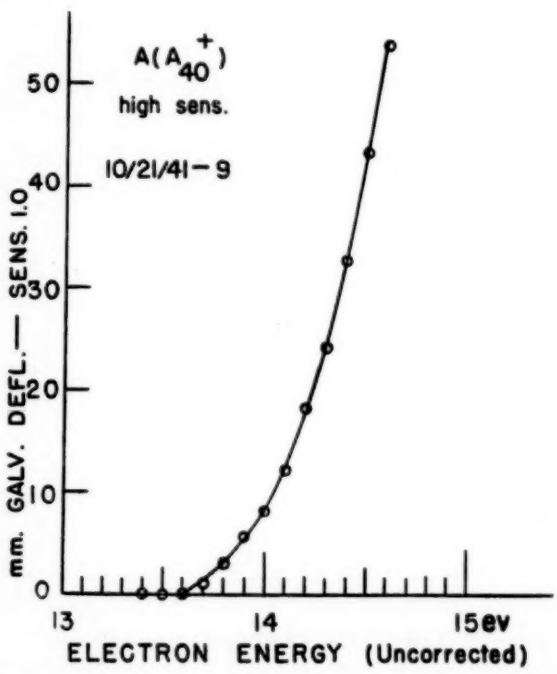


FIG. 1. Representative appearance-potential curve for the calibrating gas, argon, on high sensitivity as used for calibration.

Appearance Potential Measurements

A Barth[8] type amplifier, employing a 10^{10} ohm grid shunt, and having a sensitivity of 1.08×10^{5} mm/volt was used to measure the ion currents at electron energy intervals of 0.1 volt for the more abundant ions and at larger intervals for the less abundant ions. The zero of the amplifier was read before and after each ion current reading in order to eliminate distortion of the appearance-potential curves by amplifier drift. The voltage scale for the electrons was corrected by the difference between the known[4] and observed ionization potential of argon, which was admitted to the mass spectrometer along with the HCOOH so that the correction could be determined for each set of observations. Representative appearance-potential curves are shown

TABLE II. Energies relative to HCOOH of possible states of aggregation of 2H, 2O, and C.

Arbitrary designation i	State of aggregation	Expression for energy, E_i	E_i, e.v.
a	HCOOH	———	0.00
b	C+2H+2O	Q_f(HCOOH)+L(C) +D(O_2)+D(H_2)	18.82
c	CH+H+2O	E_b-D(CH)	15.35
d	CH_2+2O	$E_b-D(CH_2)$	11.82
e	H_2+O_2+C	$E_b-D(O_2)-D(H_2)$	9.26
f	2H+O_2+C	$E_b-D(O_2)$	13.74
g	H_2+2O+C	$E_b-D(H_2)$	14.34
h	CO_2+H_2	Q_f(HCOOH)−$Q_f(CO_2)$	−.25
i	CO_2+2H	$E_b+D(H_2)$	4.23
j	H_2O+CO	Q_f(HCOOH)−Q_f(CO) −$Q_f(H_2O)$	.18
k	OH+H+CO	$E_j+E(H_2O\rightarrow OH+H)$	5.36
l	2H+O+CO	E_b+D(OH)	9.67
m	H+OH+C+O	E_b-D(OH)	14.51
n	H_2+CO+O	E_b+D(OH)−$D(H_2)$	5.19
o	H_2O+C+O	$E_b-D(H_2O)$	9.33
p	CH+OH+O	E_b-D(OH)−D(CH)	11.04
q	2OH+C	E_b-2D(OH)	10.20
r	O_2+CH+H	$E_b-D(O_2)-D$(CH)	10.27
t	CH_2+O_2	$E_b-D(O_2)-D(CH_2)$	6.74
u	HOOH+C	Q_f(HCOOH)+L(C) −Q_f(HOOH)	7.80
v_1	HCHO+O	Q_f(HCOOH)+½$D(O_2)$ −Q_f(HCHO)	5.14
v_2	HCOH+O	approx. Q_f(HCOOH) +½$D(O_2)-Q_f(CH_3OH)$ +$E(CH_3OH\rightarrow$HCOH +H_2)	8.0 (roughly)
w_1	HCO+H+O	$E_{v1}+E$(HCHO→HCO+H)	9.62
w_2	HCO+OH	$E_{w1}-D$(OH)	5.31
w_3	COH+O+H	$E_{v2}+D$(CH) (roughly)	11.5 (roughly)
w_4	COH+OH	$E_{w2}-D$(OH)	7.2 (roughly)
x_1	HO_2+C+H	$E_u+E(H_2O_2\rightarrow H+HO_2)$ approx. $E_u+E(H_2O$ →HO+H)	12.98
x_2	HO_2+CH	$E_{x1}-D$(CH)	9.51
y	HCOO+H	approx. B(O−H) alcohols	4.6
z	H+COOH	E(HCHO→HCO+H)	4.48

[8] G. Barth, Zeits. f. Physik 87, 399 (1934); D. B. Penick, Rev. Sci. Inst. 6, 115 (1935).

TABLE III. Values of thermodynamic quantities used in calculating energies of states in Table II.

Quantity or reaction†	Energy in e.v.	Reference
Q_f(HCOOH)	3.85	a
$Q_f(H_2O_2)$	1.46	a
$Q_f(CO_2)$	4.10	a
Q_f(CO)	1.16	a
$Q_f(H_2O)$	2.51	a
Q_f(HCHO)	1.25	a
$Q_f(CH_3OH)$	2.10	a
$D(O_2)$	5.08	10
$D(H_2)$	4.48	10
D(CH)	3.47	10
$D(H_2O)$	9.49	10
D(CO)	9.14	10*
$D(CH_2)$	7.0	1
D(OH)	4.31	10
L(C)	5.41	10*
$CH_3OH \rightarrow HCOH + H_2$	3.7	1, 2
$H_2CO \rightarrow HCO + H$	4.48	11
$H_2O \rightarrow OH + H$	5.18	10
B(C=O) formaldehyde	6.16	b
B(O—O)	1.51	b
B(C=O)—B(C—O)	3.6	2
B(O—H) alcohols	4.6	c

† Q_f( ) = heat of formation of compound indicated; D( ) = heat of dissociation of compound or radical indicated; B( ) = bond energy; L(C) = heat of sublimation of carbon.

[a] B. F. Bichowsky and F. D. Rossini, *The Thermochemistry of Chemical Substances* (Reinhold Publishing Corporation, New York, 1936).

[b] L. Pauling, *Nature of the Chemical Bond* (Cornell University Press, Ithaca, New York, 1940).

[c] N. V. Sidgwick, *The Covalent Link in Chemistry* (Cornell University Press, Ithaca, New York, 1933).

* H. D. Hagstrum (Phys. Rev. 71, 376 (1947)) has recently suggested an interpretation of the band spectrum of CO which is consistent with the value D(CO) = 9.6 volts. If this should be correct, use of Herzberg's cycle (reference 10, p. 521) would yield the value L(C) = 5.87 volts. The energies of states *b*, *c*, *d*, *e*, *f*, *g*, *m*, *o*, *p*, *q*, *r*, *t*, *u*, x_1, and x_2 in Table II would thus have to be increased by 0.46 volts. This would in no way affect decisions concerning what processes are responsible for the ions observed from HCOOH; and it would affect detailed conclusions in only one process, namely that giving rise to CH_2^+. We would then obtain for this process $W(X) = 0.4 \pm 0.7$ volts, which is insignificantly more desirable than the value shown in Table I.

in Figs. 1 to 8 inclusive, in which the appropriate corrections have not been applied to the voltage scale.

The uncorrected first appearance potential of an ion was taken to be the observed electron energy at which the first perceptible ion current was detected, and the second one that at which the first perceptible "break" occurred in the appearance-potential curve.[9]

A "break" in an appearance-potential curve is a point separating a region of lower slope from one of higher slope. When the two regions are separated by a transition region, the "break" is taken to be at the beginning of the transition region. Depending on the nature of the processes involved and the conditions under which an appearance potential curve is recorded, a "break" may or may not be very pronounced. A "break" may arise from three sources, in general. First, it may result from instrumental factors such as a sudden increase in electron current. Second, it may be due to the natural trend of the appearance-potential curve for a single process, which in turn depends on the shapes and relative positions of the potential functions of the states involved. Third, it may be the nocsequence of the onset of a new process of production of the ions in question. Only in the last case should we regard the "break" as a second appearance potential. It is not difficult to take into account the instrumental factors. Thus, in the case of an abrupt change in slope, the decision concerning its significance as a second appearance potential can usually be made with confidence. When the "break" is not well defined, however, the decision cannot be made with such confidence, and

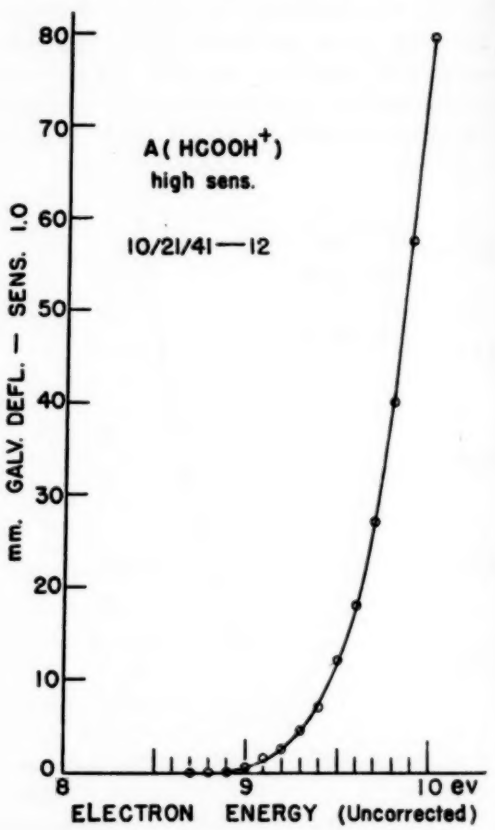


FIG. 2. Representative appearance-potential curve for $HCOOH^+$ on high sensitivity, used to determine the first ionization potential.

[9] This is the usual method of using appearance-potential curves. A different method has recently been used by R. H. Vought, Phys. Rev. 71, 93 (1947). For a discussion of the two methods see the following paper.

the personal factor may be important even in the detection of the "break." Thus, it is possible to conclude that a second appearance potential exists when only the natural trend of the ion-production efficiency of a single process is represented, or to overlook the onset of a second process when it actually occurs. In doubtful instances further experimentation may aid in reaching a conclusion; but in some cases the reasonableness of the assignment of a second appearance potential will rest solely on the reasonableness of the conclusions to which it leads, and the conclusions must then be recognized as having resulted from somewhat indecisive evidence.

The appearance potentials of ions produced in formic acid monomer by electron impact are listed in Table I, along with their relative abundances for 65-volt electrons. Except for the less abundant ions, an effort was made to adjust conditions so that each appearance potential was determined from curves in which the onset of production of ions related to the appearance potential had about the same degree of sharpness that characterized the onset of ionization in argon. The extent to which such conditions were realized had a major influence on the estimated error for the separate determinations. The listed errors are meant to refer only to the estimated precision of locating on an individual appearance-potential curve, the electron energy corresponding to the onset of ionization for a particular process, and they include the similar estimated error for the calibrating gas. Since the onset of ionization for the calibrating gas was very sharp (see Fig. 1), and since less sharp onset of ionization for any other ion corresponds to lower initial slope of the appearance-potential curve, thereby decreasing the estimated precision of locating the critical points, it is believed that the estimated error includes in it most of the potential error due to electron-energy distribution; for there would be essentially no such error if all the curves had the same initial high slope. All the first appearance potentials of ions from HCOOH for which corresponding reactions are listed in Table I are the result of at least three separate determinations (except that there were only two in the case of CH_2^+), and in all these cases, the estimated error is both greater than

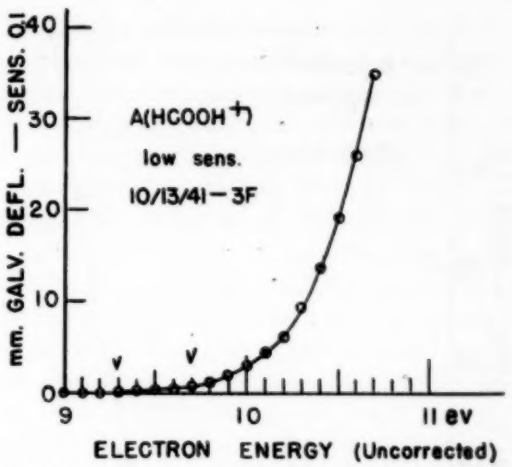


FIG. 3. Representative appearance-potential curve for $HCOOH^+$ on low sensitivity, used to determine the second ionization potential.

the maximum deviation from the mean and about twice as great as the average magnitude of the deviation from the mean.

III. DISCUSSION

In order to arrive at the interpretations of the appearance potentials indicated in Table I extensive use was made of the calculated energies (relative to HCOOH=O) of the various possible states of aggregation of 2H, 2O, and C. Table II lists the states in the order in which their energies were calculated. and gives the expressions for the energies to indicate how they were calculated. The arbitrary designations of the states are simply for convenience in writing the expressions for the energies. The values of the various thermodynamical quantities employed in reducing the energy expressions to the numerical values shown in Table II are given in Table III along with references to the sources. Quantities originally given in kcal./mol have been converted by means of the factor, 23.05 kcal./mol = 1 electron-volt.[10] Heats of formation have not been corrected to absolute zero, since an examination of some such corrections made by Burton[11] indicated that they would be of the order of magnitude of the minimum experimental uncertainties. Other data besides those appearing in

[10] G. Herzberg, *Molecular Spectra and Molecular Structure*, Vol. I, *Diatomic Molecules* (Prentice-Hall, Inc., New York, 1939), p. 4.

[11] M. Burton, J. Chem. Phys. 6, 818 (1938).

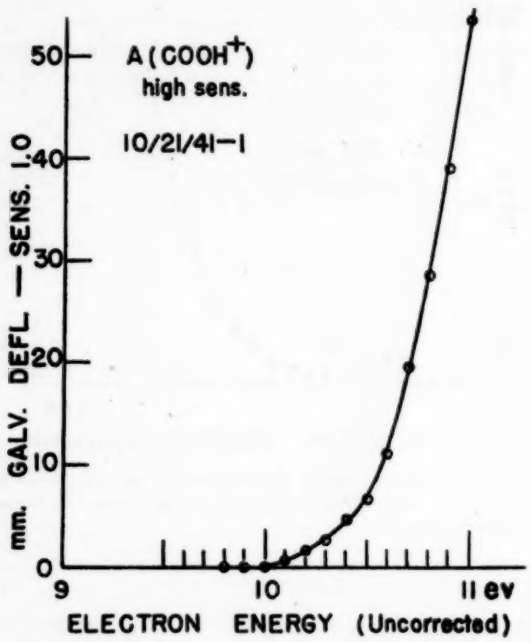


FIG. 4. Representative appearance-potential curve for $COOH^+$ on high sensitivity, used to determine the first appearance potential.

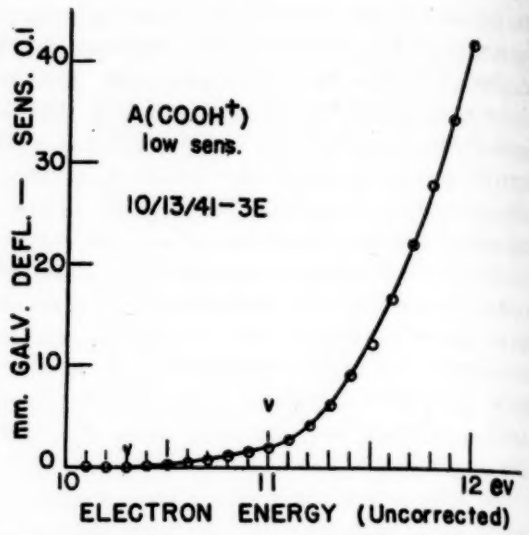


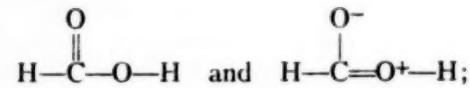
FIG. 5. Representative appearance-potential curve for $COOH^+$ on low sensitivity, used to determine the second appearance potential.

the tables will be required for a full discussion of the production of the ions, but they will be introduced in the treatment of the ions individually.

$M/e=46$

Appearance potential curves taken on high and low sensitivity, respectively (e.g. see Figs. 2 and 3), permitted the determination of two ionization potentials of HCOOH, namely at 11.0 ± 0.1 and 11.6 ± 0.2 volts. Price and Evans[12] from the ultraviolet absorption spectrum, have obtained the value 11.29 volts for the ionization potential of formic acid monomer, and have associated that value with the removal of one of the essentially non-bonding $2p$ electrons of the carboxyl oxygen. Although they do not state their experimental error, it is reasonable to assume that it is not more than a few hundredths of a volt at most.[13] Thus, there is, between the value of Price and Evans and the lower value of the present work, a discrepancy of 0.2 v. beyond the estimated experimental error of the electron impact data.

It is unlikely (although not impossible) that the electron impact data are so much in error—particularly in the indicated direction, since ionization potentials obtained by electron impact are more likely to be on the high side. Moreover, if we assume that a large error does exist and that our first ionization potential is to be identified with the spectroscopic value, then the possible explanations of the second ionization potential are not very appealing. For example, we might assume it to be due to ionization of the hydroxyl oxygen, which would probably have a higher ionization potential than the carboxyl oxygen as a result of resonance[6,14,15,16] between the structures:



but it would be difficult to explain the greater probability of ionization of the hydroxyl oxygen. (It is apparent from Fig. 3, that the mode of ionization corresponding to the second appearance potential is the more probable when the

[12] W. C. Price and W. M. Evans, Proc. Roy. Soc. A162, 110 (1937).

[13] For the ionization potential of formaldehyde obtained in a similar manner Price gives ±0.01 volts as the error; J. Chem. Phys. 3, 256 (1937).

[14] L. Pauling, *Nature of the Chemical Bond* (Cornell University Press, Ithaca, New York, 1940).

[15] R. J. B. Marsden and L. E. Sutton, J. Chem. Soc. (London) 1383 (1936).

[16] L. Pauling and J. Sherman, J. Chem. Phys. 1, 606 (1933).

electrons have enough energy to excite either mode.) Of course, we cannot categorically exclude the possibility that the second appearance potential is spurious, arising from the, effects discussed in the third paragraph of *Appearance Potential Measurements*. However, the possibility that instrumental effects are responsible is extremely unlikely; and it will soon be seen that a second appearance potential so near the first occurred only in the case of three types of ions and that one explanation seems plausible for all three.

If we discard the assumption that a large error exists in the measurement of the first ionization potential, and if we consider, rather, that our second ionization potential is to be identified with the spectroscopic value, there is a discrepancy of only 0.1 volt beyond the estimated experimental error, and it is in an understandable direction. However, we must account for the lower ionization potential at 11.0±0.1 volts. This seems readily done by assuming, as Coop, Davidson, and Sutton have suggested as a possible explanation of their dipole moment data,[17] that there exist two forms of HCOOH, one with the hydroxyl hydrogen *cis* to the carbonyl oxygen, and the other with it *trans*. Marsden and Sutton[15] have stated that the *cis*-form*** of the carboxylic esters should be the more stable; and we would expect the same to be true for HCOOH. In fact, Williams[18] has found the infra-red spectrum of HCOOH to be compatible with the *cis*-form; but, in conversation with one of us, has stated that his data do not preclude the possibility of existence of a *trans*-form in low concentration. The *trans*-form, then, should appear in lower concentration than the *cis*, if at all. Moreover, the ionization potential of the carbonyl oxygen should be lower in the *trans*-form than in the *cis*; for, in the latter, the greater proximity of the hydroxyl hydrogen provides a greater attractive force on the $2p$ electrons of the carboxyl oxygen. These considerations are consistent with Fig. 3, in which it is evident that the lower ionization potential is associated with the production of ions in low abundance relative to the ions associated with the higher ionization potential. The implication of the apparent relative abundance of *cis*- and *trans*-forms will be presented at the end of the paper, since the estimate of the abundance will involve other ions as well as HCOOH.

Before deciding that the assumption of a *trans*-form of HCOOH is plausible, it seems necessary at least to correlate the structure of the molecule with the difference between the ionization potentials 11.0±0.1 volts and 11.6 ±0.2 volts which we would assume to apply to the $2p$ electrons of the carbonyl oxygen in the *trans*- and *cis*-forms, respectively. Let us assume the potential field for these electrons to be divided into two parts: V_2 due to the hydroxyl hydrogen, and V_1 due to the rest of the molecule

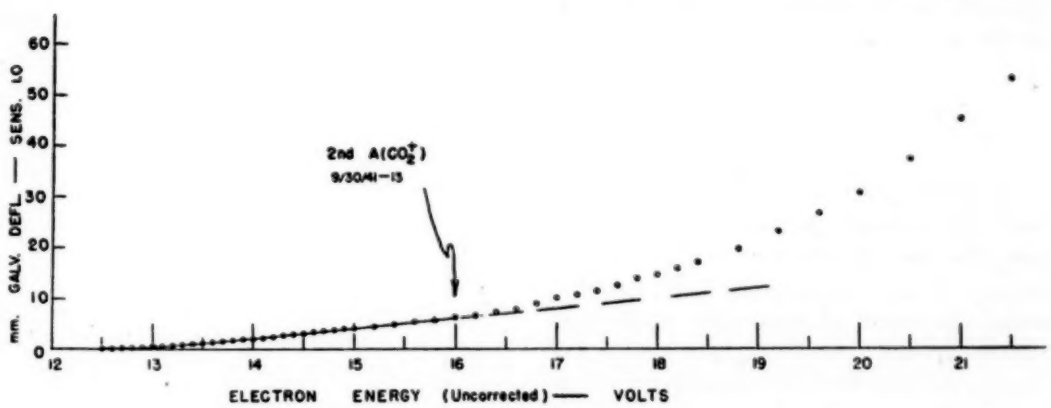


FIG. 6. Appearance potential-curve for CO_2^+ on low sensitivity, used to determine the second appearance potential.

[17] I. E. Coop, N. R. Davidson, and L. E. Sutton, J. Chem. Phys. 6, 905 (1938).

*** They use "*trans*"; but they have labelled the forms in the opposite sense, referring them to the alkyl radical on the carbon atom.

[18] V. Z. Williams, J. Chem. Phys. 15, 232 (1947).

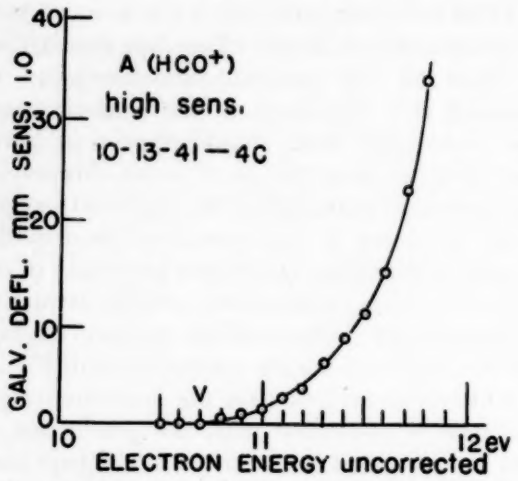


FIG. 7. Representative appearance-potential curve for HCO^+ on high sensitivity, used to determine the first appearance potential.

Let us assume further that the *trans*-form differs from the *cis*-form described by Williams[18] only in that the hydroxyl hydrogen is rotated 180° around the C—O bond as axis. To a first approximation, V_1 will be the same for both forms, and we need consider only the field due to the hydroxyl hydrogen in computing the difference in ionization potentials of the two forms. If the hydroxyl hydrogen is considered to have an effective charge e_h'(esu), we may approximate the difference ΔIP between the ionization potentials by the difference between the coulomb interaction energies (between the hydroxyl hydrogen and a $2p$ electron of the carbonyl oxygen) for the two forms; or:

$$\Delta IP = -e_h' e\left\{\frac{1}{r_{cis}} - \frac{1}{r_{trans}}\right\},$$

where $e = -4.80\times10^{-10}$ esu, and where r_{cis} and r_{trans} are the distance between the hydroxyl hydrogen and the $2p$ electron in the *cis*- and *trans*-forms. If we accept the corresponding interatomic distances as reasonable approximations, we have for r_{cis} and r_{trans} the values 2.39×10^{-8} cm and 3.08×10^{-8} cm, respectively. These lead to:

$$\Delta IP = e_h'\times4.47\times10^{-3} \text{ ergs.}$$

We obtain a more convenient form by expressing ΔIP in electron-volts and the residual charge, e_h, in terms of fractions of a positive electronic charge. We then have:

$$\Delta IP = e_h\times1.34 \text{ ev.}$$

The observed difference in ionization potentials (0.6 ± 0.3 ev) then leads to $e_h = 0.45\pm0.23$ positive electron unit. This should correspond to the value assigned from dipole moment data, which, according to Sidgwick (reference c, Table III), is 0.31 positive electron unit. Thus, the observed ΔIP is not of the wrong order of magnitude to be interpreted in terms of *cis*- and *trans*-forms of HCOOH.

For the rest of the discussion, then, it will be assumed that 11.0 ± 0.1 volts is the ionization potential corresponding to the removal of a $2p$ electron from the carbonyl oxygen in molecules of the *trans*-form, and that 11.6 ± 0.2 volts is the corresponding value for molecules of the *cis*-form. This assumption will not affect the selection of probable dissociation processes; but it will be useful in interpreting double appearance potentials in the case of some other ions.

$M/e = 45$

The ion corresponding to the removal of a hydrogen atom from HCOOH has appearance potentials at 12.1 ± 0.1 volts and 12.9 ± 0.2 volts, determined from appearance-potential curves similar to those shown in Figs. 4 and 5. The difference between the appearance potentials is essentially the same as found for $HCOOH^+$.

If we assume that this ion is formed simply by the rupture of the C—H or O—H bond, accompanied by ionization of the radical, the calculated appearance potential (referring to

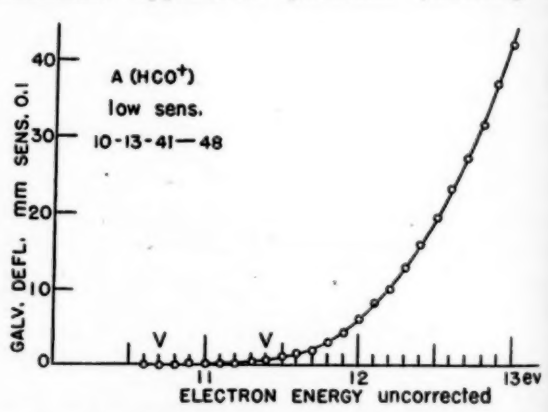


FIG. 8. Representative appearance-potential curve for HCO^+ on low sensitivity, used to determine the second appearance potential.

states of aggregation y and z) is:

either (for y)

$$4.6+I(\mathrm{HCOO})+W(X)=A(\mathrm{HCOO^+}),$$

or (for z)

$$4.5+I(\mathrm{COOH})+W(X)=A(\mathrm{COOH^+}).$$

From the observed appearance potentials we then calculate:

$$I_1(\mathrm{HCOO})\leqq 7.5\pm 0.1 \text{ v};$$
$$I_2(\mathrm{HCOO})\leqq 8.3\pm 0.2 \text{ v}$$

and

$$I_1(\mathrm{COOH})\leqq 7.6\pm 0.1 \text{ v};$$
$$I_2(\mathrm{COOH})\leqq 8.4\pm 0.2 \text{ v},$$

which are unreasonably low.

We find much more plausible energy relationships if we assume a process similar to that found convenient by Cummings and Bleakney[2] in explaining some of the appearance potentials in methyl and ethyl alcohol. If the ionization of the radical takes place at the carbonyl oxygen, the ion can have the form $\mathrm{H-O-C\equiv O^+}$ (utilizing the trivalency of O^+), which would require for its production an energy less than that for state z by the difference between the triple and double bond energies. We have taken for this difference: $D(\mathrm{CO})-B(\mathrm{C=O})$, or 3.0 v. The calculated appearance potential is, then:

$$1.5+I(\mathrm{COOH})+W(X)=A(\mathrm{COOH^+}),$$

which by use of the observed appearance potentials gives

$$I_1(\mathrm{COOH})\leqq 10.6\pm 0.1 \text{ v}; \quad I_2\leqq 11.4\pm 0.2 \text{ v}.$$

Considering the crudeness of the approximation used in obtaining the dissociation energy, these values are in good agreement with the values for the ionization potential of the carbonyl oxygen (assumed) in the molecule, and it seems plausible to assume that the two appearance potentials in the present case arise from the *trans-* and *cis-*forms of HCOOH just as we did in the case of the molecular ion.

Since the utilization of the trivalency of O^+ is necessary to explain the low appearance potential of the ion of $M/e=45$, it must be the C–H hydrogen which is rejected in the dissociation; otherwise there would not be enough free valencies on the carbon atom to permit the formation of the triple bond with the oxygen. The trivalency of O^+ could be used in the alternate form of the ion in which the ionization occurs at the hydroxyl oxygen: $\mathrm{H-O^+=C=O}$, and consideration of appearance potentials alone does not permit excluding this possibility. However, since, as was pointed out earlier, the ionization potential of the hydroxyl oxygen is probably higher than that of the carbonyl oxygen, the ionization probably takes place at the carbonyl oxygen. In addition, if it is the hydroxyl oxygen which is ionized in the ion now under discussion, the occurrence of two appearance potentials is not easily explained. Finally, it will be seen later that independent arguments lead to the conclusion that it is probably the carbonyl oxygen which is ionized in the case of the ion of $M/e=29$.

$M/e=44$

Curves taken on high sensitivity yielded the value 14.3 ± 0.1 v for the first appearance potential of CO_2^+; and curves taken on low sensitivity, such as that shown in Fig. 6, gave the value 18 ± 1.0 v for the second appearance potential. CO_2^+ can be produced from HCOOH only in the two reactions indicated in Table I; and the assignment of the first appearance potential must be correct, for it corresponds to an electron energy too low to cause a higher degree of dissociation than is indicated. If the second appearance potential is real, as it seems to be, it can be associated only with the other possible reaction. If the appearance potential curve has been properly interpreted, it seems that the reaction giving the higher degree of dissociation is the more probable one at high electron energies.

In calculating the energies required for the two reactions, $I(CO_2)$ was taken to be 13.73 ± 0.01 v as given by Price and Simpson.[19] It should be noted that $W(X)$ is small for both reactions.

$M/e=34$: and $M/e=33$

As is indicated in Table I, these ions appeared in very low abundance relative to the molecular ion, $M/e=46$, and are attributed to a slight

[19] W. C. Price and D. M. Simpson, Proc. Roy. Soc. **A169**, 501 (1939).

impurity of H_2O_2 in the sample. The complete reasoning on which this conclusion is based is rather extended; and, since the final conclusion contributes nothing concerning the molecule being studied, it seems hardly worth while taking space for the argument.

$M/e=32$

These ions had the same appearance potential in background and sample; and no increment to the background peak could be observed when the sample was admitted. Since the background at $M/e=32$ was only 0.25 percent of the intensity of the peak at $M/e=46$ in the sample, it is evident that very few, if any, O_2^+ ions were formed from HCOOH. If it is assumed that the background ions are O_2^+, we deduce the value $I(O_2)=13.2\pm0.6$ v, which is to be compared with 12.3 ± 0.1 v given by Hagstrum and Tate.[20] The discrepancy of 0.2 v beyond the combined experimental errors may probably be taken as indicative of the specific error introduced by comparing an appearance-potential curve of very low slope (as was the case for $M/e=32$) with one of very high slope (as was the case for the calibrating gas).

$M/e=30$

Owing to the isotope effect, there should be ions of $M/e=30$ in the amount of 1.16 percent of mass 29 (or 2.15 percent of mass 46) plus 0.2 percent of mass 28 (or 0.03 percent of mass 46), making a total of 2.18 percent of mass 46.[a] The observed number, 2.64 percent of HCOOH, was slightly greater. The difference might be due to an abnormal distribution of isotopes or to a few H_2CO^+ ions. The calculated appearance potential in the process (cf. state v_1):

$$HCOOH+A(H_2CO^+)\rightarrow H_2CO^+ + C + W(X) + e$$

is 15.97 volts (using Price's value[14] for $I(H_2CO)$), which is higher than the observed 14.1 ± 0.3 v. It will presently be seen that ions of $M/e=29$ are produced with two appearance potentials, the lower of which is associated with the production of ions in rather small quantities. The appearance potentials of the mass 30 ions, then, is certainly to be identified with the higher appearance potential (13.2 ± 0.3 v) of the mass 29 ions. There is a discrepancy of 0.3 volt beyond the combined limits of error of measurement. This is the same order of magnitude as the discrepancy found in the case of the mass 32 ions, which is not surprising considering the fact that the appearance-potential curves for 30 and 32 had nearly the same low slope.

$M/e=29$

Ions of $M/e=29$ were the most abundant of the ions produced from HCOOH by electron impact. Figures 7 and 8 are characteristic of those taken at high and low sensitivities to determine, respectively, a lower appearance potential at 12.8 ± 0.2 v and a higher appearance potential at 13.2 ± 0.3 v. The difference is essentially the same as found in the case of mass 46 and mass 45.

Of the modes of straightforward dissociation of HCOOH giving rise to a radical of mass 29, the one yielding the final state of aggregation w_2 requires the least energy. The calculated appearance potential of HCO^+ in the corresponding reaction,

$$HCOOH+A(HCO^+)\rightarrow HCO^+ + OH + W(X) + e,$$

is $I(HCO)+5.3$ v. Using the two observed appearance potentials we obtain $I_1(HCO)\leqq 7.5\pm0.2$ v and $I_2(HCO)\leqq 7.9\pm0.3$ v. These seem too low, since we expect $I(HCO)$ to be nearly the same as $I(HCOOH)$, for which we have obtained the values 11.0 ± 0.1 v and 11.6 ± 0.2 v ascribed, respectively, to the (presumed) *trans*- and (known) *cis*-forms of HCOOH. The energy relationships are much better if we assume, in the reaction above, that the HCO^+ ion has the structure $H-C\equiv O^+$. In this case the dissociation energy is reduced by about 3.9 volts as in the analagous situation considered for mass 45; and we obtain

$$I_1(HCO)\leqq 10.5\pm0.2 \text{ v}$$

and

$$I_2(HCO)\leqq 10.9\pm0.3 \text{ v}.$$

These are in good agreement with the values

$$I_1(COOH)\leqq 10.6\pm0.1 \text{ v}$$

[20] H. D. Hagstrum and J. T. Tate, Phys. Rev. 59, 354 (1934).

[a] On the basis of relative abundances of C, H, O, isotopes as given by G. T. Seaborg, Rev. Mod. Phys. 16, 1 (1944).

and

$$I_2(\text{COOH}) \leqq 11.4 \pm 0.2 \text{ v}$$

obtained in the discussion of mass 45; and, considering the uncertainty in the approximation used for difference in energy between $-C{=}O^+$ and $-C{\equiv}O^+$, they are in reasonable agreement with the values obtained for $I(\text{HCOOH})$.

In the case of HCO^+, then, it appears that the hydroxyl group is rejected in the dissociation of HCOOH and that the carbonyl bond in the ion is a triple bond, made possible only by ionization of the carbonyl oxygen. The two appearance potentials are consistent with the assumption that a *trans*-form of HCOOH exists.

$M/e=28$

Some of the appearance-potential curves for these ions showed a second appearance potential at about 16 volts; but this was shown to be due to $I(N_2)$ arising from a temporary leak.

The CO^+ ions evidently arise from the reaction (cf. state j),

$$\text{HCOOH}+A(\text{CO}^+) \rightarrow \text{CO}^+ + \text{H}_2\text{O} + W(X) + e.$$

The energy required is $I(\text{CO})+0.2$ v. Using Vaughan's value[21] $I(\text{CO})=13.9\pm0.2$ v, the calculated appearance potential is 14.1 ± 0.2 v, in fair agreement with the observed value of 14.6 ± 0.3 v. The excess energy in the reaction is 0.5 ± 0.5 v. Other reactions in which CO^+ might be produced (cf. states k, n) would require an energy about 5 volts greater than the observed appearance potential.

$M/e=18$

H_2O^+ ions appeared at 13.1 ± 0.1 volts electron energy, with no second appearance potential up to 25 volts. The reaction responsible for these ions must be:

$$\text{HCOOH}+A(\text{H}_2\text{O}^+) \rightarrow \text{H}_2\text{O}^+ + \text{CO} + W(X) + e.$$

The minimum required energy is $I(\text{H}_2\text{O})+0.18$ v, which, on the basis of $I(\text{H}_2\text{O})=12.59\pm0.05$ v,[22] gives 12.77 ± 0.05 v in good agreement with the observed value. The excess energy of the reaction is 0.3 ± 0.15v.

It is interesting to note that the assigned reaction for the production of H_2O^+ involves the same dissociation process as that for CO^+ and that the dissociation energy is so small that the ions have appearance potentials not much different from their ionization potentials. This suggests that H_2O^+ and CO^+ may have occurred simply by ionization of the product molecules in the possible thermal decomposition,

$$\text{HCOOH} \rightarrow \text{H}_2\text{O} + \text{CO},$$

at the filament or in the ionization chamber. However, the ions were much too abundant to be accounted for in this way. Since differential pumping was employed, there is little possibility that products of dissociation at the filament could have entered the ionization chamber. Moreover, as remarked under the discussion of the sample, Bonner and Hofstadter found no perceptible decomposition of formic acid vapor held for 36 hours at 150°C, which is probably only slightly lower than the temperature of the ionization chamber. Unless there was some strong catalytic effect favoring thermal decomposition in the present case, it seems most likely that the ions arose as the result of electron impact according to the assigned reactions.

$M/e=17$

There are two energetically indistinguishable reactions which are consistent with the observed appearance potential. Corresponding to states of aggregation designated as w_2 and k, these are, respectively,

$$\text{HCOOH}+A(\text{OH}^+) \rightarrow \text{OH}^+ + \text{HCO} + W(X) + e$$

and

$$\text{HCOOH}+A(\text{OH}^+) \rightarrow \text{OH}^+ + \text{H} + \text{CO} + W(X) + e.$$

Using $I(\text{OH}) \leqq 13.6\pm0.2$ as given by Mann *et al*,[23] and 5.3 and 5.4 volts as the dissociation energies, we calculate the expected appearance potentials 18.9 ± 0.2 v and 19.0 ± 0.2 v for the two reactions, leading to $W(X) \geqq 0.3\pm0.4$ v and $\geqq 0.2\pm0.4$ v. Other reactions in which OH^+ might be produced require appearance potentials at least 4.6 volts greater than that observed.

$M/e=16$

For O^+ ions there was only one acceptable measurement of the appearance potential not

[21] A. L. Vaughan, Phys. Rev. **38**, 1687 (1931).
[22] L. G. Smith and W. Bleakney, Phys. Rev. **49**, 883A (1936).
[23] M. M. Mann, A. Hustrulid, and J. J. Tate, Phys. Rev. **58**, 340 (1940).

interfered with by background. This gave 21.0 ±0.6 volt for the appearance potential. Bacher and Goudsmidt[7] give 13.55 v for $I(O)$, so the process giving rise to O^+ must be one in which the dissociation energy is equal to or less than 7.4±0.6 v. Of the thirteen modes of dissociation which might give rise to an oxygen atom, there are three with low enough dissociation energies not to be excluded as possibilities. These correspond to final states of aggregation and dissociation energies as follows:

v_1	H_2CO+O	5.14 volts,
v_2	$HCOH+O$	8.0 volts,
r	H_2+CO+O	5.19 volts.

$M/e=14$

There are two processes in which CH_2^+ ions might arise:

$$HCOOH+A(CH_2^+)\rightarrow CH_2^++O_2+W(X)+e$$

and

$$HCOOH+A(CH_2^+)\rightarrow CH_2^++2O+W(X)+e.$$

On the basis of the dissociation energies for states t and d and Smith's[1] value $I(CH_2)\leqq 12.0\pm 0.2$ v the calculated appearance potentials are, respectively, $\leqq 18.7\pm 0.2$ v and $\leqq 23.8\pm 0.2$ v. If the first reaction is the one occurring, $W(X)\geqq 6.0\pm 0.7$ v; if the second one occurs, $W(X)\geqq 0.9\pm 0.7$ v. It seems more likely, then, that the second reaction is the one which produces the observed ions.

$M/e=13$

There are four reactions in which CH^+ ions might be produced:

$HCOOH+A(CH^+)\rightarrow$
$\quad CH^++HO_2+W(X)+e$ (cf. state x_1),
$HCOOH+A(CH^+)\rightarrow$
$\quad CH^++O_2+H+W(X)+e$ (cf. state r),
$HCOOH+A(CH^+)\rightarrow$
$\quad CH^++OH+O+W(X)+e$ (cf. state p),
$HCOOH+A(CH^+)\rightarrow$
$\quad CH^++2O+H+W(X)+e$ (cf. state c).

The corresponding minimum required energies are:

$I(CH)+9.5$ volts, $I(CH)+10.2$ volts,
$I(CH)+11.0$ volts, $I(CH)+15.4$ volts.

It is unlikely that the last reaction is responsible for the second appearance potential. If it were, then we should deduce $I(CH)\leqq 12.0$ v; and, following the reasoning of Kusch, Hustrulid, and Tate,[24] we should calculate the rather improbable value $W(X)\geqq 3.4\pm 1.1$ v in their reaction $HCN\rightarrow CH^++N$. Therefore, either the second or third reaction is probably responsible for the second appearance potential. These choices give, respectively, $I(CH)\leqq 17.2\pm 0.6$ v and $I(CH)\leqq 16.4\pm 0.6$ v; and since the authors just mentioned give $I(CH)\leqq 15.4\pm 0.5$ v, we calculate $W(X)\geqq 1.8\pm 1.1$ v and $W(X)\geqq 1.0\pm 1.1$ v, respectively, for the two reactions. Neither of these values is unreasonable, and neither reaction can be excluded.

Since either the second or third reaction may be responsible for the second appearance potential, either the first or second reaction may be responsible for the first appearance potential. Considering these, we deduce $I(CH)\leqq 15.9\pm 0.5$ v and $I(CH)\leqq 15.2\pm 0.5$ v, or using $I(CH)\leqq 15.4\pm 0.5$ v, as above, we deduce $W(X)\geqq 0.5\pm 1.0$ v and $W(X)\geqq -0.2\pm 1.0$ v.

$M/e=12$

Any one of four reactions may account for the first appearance potential at 23.9±0.2 v. These correspond to the following states of aggregation of products, calculated minimum appearance potentials, and calculated excess energies:

	State	$A(C)_{min}$ (volts)	$W(X)$ (volts)
e	$H_2+O_2+C^+$	$9.26+I(C)$	3.4±0.2
o	H_2O+C^++O	$9.3\ +I(C)$	3.4±0.2
q	$2OH+C^+$	$10.2\ +I(C)$	2.5±0.2
x_1	HO_2+C^++H	$13.0+I(C)$	−0.3±0.2

Likewise, there are at least four possibilities for the second appearance potential at 29.5±1.0 v:

	State	$A(C)_{min}$ (volts)	$W(X)$ (volts)
g	H_2+2O+C^+	$14.2+I(C)$	4.0±1.0
f	$2H+O_2+C^+$	$13.7+I(C)$	4.6±1.0
m	$H+OH+C^++O$	$14.5+I(C)$	3.8±1.0
b	$2H+2O+C^+$	$18.8+I(C)$	−0.5±1.0

There are no strong arguments on the basis of which a more restricted assignment can be made.

[24] P. Kusch, A. Hustrulid, and J. T. Tate, Phys. Rev. 52, 843 (1937).

$M/e=2$

H_2^+ ions occurred in very low abundance, about 0.03 percent of mass 46 for 65-volt electrons. Their appearance potential, 16.0±0.3 v, lay in the range of that of the background ions, 16.6±0.5 v; but it is quite certain that the observed appearance potential is properly assigned to ions from HCOOH. The estimated error may be too small, however, both because of the low slope of the appearance-potential curves for these ions and because of the uncertainty introduced by applying, at low magnetic fields, the electron energy calibration obtained at high magnetic fields. If there had been time to do so, it would have been desirable to use molecular hydrogen for calibrating at low magnetic fields.

Unless the combined errors from all sources amount to several volts (which is unlikely), the only reaction by which the H_2^+ ions could have been produced from HCOOH is

$$HCOOH+A(H_2^+)\rightarrow H_2^++CO_2+W(X)+e.$$

On the basis of the energy of the state h (−0.25 v) and the value $I(H_2)=15.37$ v,[22] the calculated minimum appearance potential is 15.1 v, indicating an excess energy 0.9±0.3 v.

$M/e=1$

The H^+ ions were definitely produced from HCOOH. The appearance potential and the value $I(H)=13.5$ v[7] require that the reaction involve not more than 5.5±0.4 volts dissociation energy. Four possible states of aggregation and their corresponding dissociation energies are:

	State	$P(X)$ (volts)
i	CO_2+H+H	4.23
h	$OH+CO+H$	5.4
g	$HCOO+H$	4.6
z	$COOH+H$	4.5

Any one is plausible.

Negative Ions

The only negative ions found were very scarce, about 0.03 percent of $HCOOH^+$. They were definitely produced from HCOOH, since they disappeared on pumping out the sample, and they were tentatively identified as OH^- ions; but, since their appearance potential could be determined only to within ±5.0 v, it did not seem worth while to investigate them further.

IV. CONCLUSION

The electron-impact experiments on formic acid monomer required consideration of ions at one negative and seventeen positive values of M/e. The negative ions were tentatively identified as OH^- ions, but they were too rare to encourage a thorough investigation of them. Ions at $M/e=32$ were shown to be due to background, and ions at $M/e=33$, 34 were attributed to an impurity in the sample. The ions at $M/e=30$ were shown to be due almost entirely to isotope effect.

Of the positive ions arising from HCOOH as primary products of electron impact, those at $M/e=16$, 13, 12 and 1 could not be accounted for by any very restricted choice of possible reactions. Those at $M/e=17$ might have arisen from either of two possible reactions requiring nearly the same low dissociation energy and leading to calculated excess energy less than 0.5 v. Those at $M/e=14$ might have arisen from either of two reactions, but the reaction requiring the least dissociation energy involves so much excess energy as to make it unlikely, so that the ions probably arose from the reaction requiring the larger dissociation energy.

The first appearance potentials of the remaining fragment ions, namely those at $M/e=45$, 44, 29, 28, 18, and 2, were uniquely ascribable to origin of the ions in the reactions involving the least dissociation energy for the production of the corresponding radicals from HCOOH. The low appearance potentials of two of these ($M/e=45$, 29) led to the conclusion that the structure of these particular ions involved utilization of the trivalency of O^+, thus substantiating the conclusions of Cummings and Bleakney[2] regarding the production of some of the oxygen-bearing ions in methyl and ethyl alcohol. The first appearance potentials of the other four ions and the second appearance potential of CO_2^+, besides permitting determination of the five reactions involved, also permitted calculation of the excess energies required for the reactions. All of these excess energies were found to be less than 1.0 volt, thus substantiating the

view that excess energies are usually low compared to bond energies.

Two not widely separated ionization potentials were observed for HCOOH; and, in addition, a second appearance potential close to the first was observed for both $COOH^+$ and HCO^+. The postulate of a *trans*-form of HCOOH was introduced to explain the fact that the electron-impact experiments yielded a value for the first ionization potential lower than that given by Price and Evans by 0.2 volt beyond the estimated experimental error, whereas their value is more consistent with a second electron-impact ionization potential. It was shown that the difference between the two ionization potentials, determined in the electron-impact experiments, is of the right order of magnitude to be explicable in terms of *trans*- and *cis*-forms of HCOOH, and it was shown that the concept of a *trans*-form of HCOOH is useful in explaining the two appearance potentials observed for both $COOH^+$ and HCO^+.

The appearance-potential curves for $HCOOH^+$, $COOH^+$, and HCO^+ might have been explicable in other ways if there had not been available the accurate spectroscopic value for the ionization potential of the molecule, and the postulate of a *trans*-form is therefore presented with caution. Perhaps the chief reason for hesitancy is that the appearance-potential curve for argon, Fig. 1, shows some of the "tailing-off" characteristic that is found in the appearance-potential curve for $HCOOH^+$, Fig. 2. In comparing these curves, however, it will be noted that the "tailing-off" is less pronounced for argon than for $HCOOH^+$.††† Also, some of the "tailing-off" in argon is probably due to transitions to the $^2P_{\frac{1}{2}}^0$ state of A^+ which lies 0.178 v above the lowest state of A^+. Purely circumstantial evidence that the proper interpretation of the formic acid data has been made is the fact that first and second appearance potentials were assigned in the case of $HCOOH^+$, $COOH^+$, and HCO^+ before knowledge of the work of Price and Evans was obtained, and when it was known that the explanation of second appearance potentials so close to the first would present some problems; whereas there was no tendency to consider that the appearance-potential curve of argon had a "break" in it. Thus, although further work would be desirable to permit a more certain demonstration that a *trans*-form of HCOOH exists, and although we cannot suggest that the existence of a *trans*-form has been proved, certainly the electron impact data are best interpreted in terms of such a postulate.

If the assumption of the existence of a *trans*-form of HCOOH is correct, the appearance-potential curves taken on low sensitivity for $HCOOH^+$, $COOH^+$, and HCO^+ can be used to get a rough estimate of its concentration relative to the *cis*-form. This has been done by comparing the ion current at the second appearance potential with that obtained at an electron energy greater than the second appearance potential by an amount equal to the difference between the two appearance potentials. Considering the curves for the three types of ions which can be treated, we obtain for C_t/C_c, the ratio of the concentration of molecules in the *trans*-form to the concentration of molecules in the *cis*-form, the rough average 0.12 and the rough probable limits 0.05 and 0.20. It would be interesting to learn whether Price and Evans' spectrograms could have shown the ionization series for molecules in this abundance.

If it is assumed that equilibrium was attained at the temperature of the ionization chamber, the relationship

$$\ln_e(C_t/C_c) = -\Delta E/kT$$

can be used to calculate the energy difference, ΔE (ergs/molecule), between the two forms. Using for the temperature the probable value 423°K and the limits 373°K and 500°K, the concentration ratios cited above give for ΔE the rough value 1.8 kcal/mole with probable limits 1.2 and 3.0 kcal/mole.

ACKNOWLEDGMENTS

We are grateful to our associates for interest and encouragment which provided a helpful stimulus in bringing this work to completion. One of us (T.M.) is indebted to the American Cyanamid Company for its courtesy in facilitating his part in the preparation of the paper.

††† The same remarks apply to a comparison of the appearance-potential curves taken on low sensitivity; but the low sensitivity curve for argon has not been reproduced here since the scale of reproduction would not show the differences which are apparent when the full-sized original curves are superposed.

A Note on the Significance of Appearance Potentials of Ions in a Mass Spectrometer

THOMAS MARINER
Research Laboratories of the American Cyanamid Company, Stamford, Connecticut

AND

WALKER BLEAKNEY
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received July 18, 1947)

The significance of appearance potentials determined by the two common methods is treated with the aid of some simplifying assumptions. Arguments favoring the general use of the more commonly used method are given, and it is pointed out that the initial part of the appearance potential curve of an ion cannot generally lead to unambiguous conclusions concerning the nature of the ionic state involved in the production of the ion.

I. INTRODUCTION

TWO methods have been used for determining appearance potentials of ions from their appearance-potential curves (plots of ion current *versus* energy of the ionizing electrons) obtained on the mass spectrometer or simpler critical potential apparatus. One method takes as the appearance potential of an ion the corrected electron energy corresponding to the first detectable ion current; for want of a better name, and for the sake of brevity, this will be called the vanishing-current or VC method. The other method, which will be referred to as the linear-extrapolation or LE method, takes as the appearance potential of an ion the corrected electron energy corresponding to the energy intercept of the extrapolated linear part of the appearance-potential curve. Both methods have been used since early in the history of electron impact studies;* but use of the LE method has been relatively rare. In a recent paper,[1] however, the LE method has been used again, and the interesting observation has been made that its application to a fragment ion of a molecule gives, under certain conditions, quite closely the energy required for a transition through the center of the Franck-Condon region. It has also been implied that the LE method gives energies standing in a more definite relationship with the physical quantities involved than does the VC method. Since we are not in full agreement with this implication, nor with the statements of the reference paper with regard to the principal limitation of the VC method, nor with the statement of conditions under which the LE method gives quite closely the transition energy in the center of the Franck-Condon region, and since we feel that there are arguments favoring the general use of the VC method, we should like to present our views on the significance of appearance potentials obtained by the two methods. None of our remarks is to be construed as criticism of the treatment of dissociation processes in $SiCl_4$ discussed in the reference paper.

We have chosen to use a mathematical symbolism in presenting the following discussion, since it seems to be the briefest and clearest way of emphasizing the relationships between the various factors which are known to influence appearance-potential measurements. We are well aware of the departure from quantitative rigorousness which results both from over-simplification of assumptions concerning the nature of the ionization process itself and from complete neglect of the effect of variable instrumental efficiency;[2] and we do not wish to imply that the derived expressions will be generally useful in a quantitative way. However, by giving a direct and easily understandable approach to the qualitative relationships which must be taken

* For some early examples of the VC method, see Hogness and Lunn, Phys. Rev. **26**, 44 (1925); H. D. Smyth, Phys. Rev. **25**, 452 (1925); W. Bleakney, Phys. Rev. **36**, 1303 (1930); and for some early examples of the LE method see, P. T. Smith, Phys. Rev. **36**, 1293 (1930); P. T. Smith and J. T. Tate, Phys. Rev. **39**, 27 (1932); Tate, Smith, and Vaughan, Phys. Rev. **48**, 525 (1935).

[1] R. H. Vought, Phys. Rev. **71**, 93 (1947).

[2] H. D. Hagstrum and J. T. Tate, Phys. Rev. **59**, 354 (1941).

into account in formulating decisions concerning the best method of handling electron impact data, we believe that the simplified treatment is adequately justified.

II. DISCUSSION

General Significance of VC Appearance Potentials

In application of the VC method appearance-potential curves are plotted both for the ion whose appearance-potential is to be determined and for the ion whose appearance potential is known (usually the ion of a monatomic gas, such as argon, the ionization potential of which has been determined spectroscopically). For both appearance-potential curves the nominal electron energy corresponding to the first detectable ion current is determined. The difference between the nominal energy for the experimental ion and that for the calibrating ion is added to the known appearance potential of the calibrating ion to give the corrected appearance potential of the experimental ion.

It is not possible to give a precise answer to the question of the significance of an appearance potential obtained by the VC method; but general considerations lead to an interpretation which is not inconsistent with the generally satisfactory results which have been obtained by use of the method.[2–4] In the case of the calibrating atomic gas (let us consider argon, to be specific), and under conditions of reasonably high sensitivity, the first detectable ion current probably appears when the most energetic group of electrons of significant intensity in the Maxwellian distribution have just enough energy to cause a transition to the $^2P_{1\frac{1}{2}}{}^0$ state of A^+. The actual trend of the ionization curve after the onset of ionization will (apart from the unknown factors in the trend of the ionization cross section) depend on both the influence of the low lying $^2P_{\frac{1}{2}}{}^0$ state of A^+ and the finite spread of electron energies;[4] but we are concerned only with the region very close to onset of ionization. It may be that the curve should approach the energy axis asymptotically; but with sufficiently high sensitivity it appears to have a reasonably large slope even over the first 0.2 volt after the apparent onset of ionization, which it seems to be possible to locate to within about 0.05 volt (see Fig. 1 of the preceding paper). Thus, under conditions of high sensitivity, the part of the curve which is important for determination of the VC appearance potential is probably produced by electrons in a fairly narrow energy band, namely, the most energetic group of electrons which have significant intensity in the energy distribution. In the case of an ion produced from a molecule, the trend of the appearance-potential curve after the onset of ionization will be influenced by the spread of electron energies, by the distribution of vibrational levels in the ionized state, by the shape and position of the potential function of the ionized state relative to the normal state, by the existence of neighboring states involving the ion in question, and by the shape of the expectation function of the ground level, in addition to unknown factors. However, when the sensitivity is so high that the appearance-potential curve has a reasonably high slope over the first few tenths of a volt after the apparent onset of ionization (see Fig. 4 of the preceding paper), it seems reasonable to believe that the part of the curve in which we are interested is due to the same narrow energy band of electrons considered for the atomic ion, and that the first detectable ion current appears when the energy of this group of electrons is just sufficient to effect the least energetic transition (between the ground level and the ionic state in question) among the possible vertical transitions which correspond to significant values of the expectation function of the ground level. Since the same group of electrons in a narrow band at the high energy end of the electron energy spectrum is probably responsible for onset of ionization in both the calibrating gas and the experimental gas, it seems that the finite spread in electron energies should have practically no influence on VC appearance potentials when reasonably high sensitivity is used for both the calibrating ion and the experimental ions. At lower sensitivities a broader band of electron energies is effective at the apparent onset of ionization; but the effect of this broader band is probably largely accounted for in the increased estimate of experi-

[3] For example, see D. P. Stevenson and J. A. Hipple, Phys. Rev. 62, 237 (1942).
[4] W. Bleakney, Phys. Rev. 40, 496 (1932).

mental error which follows as a consequence of poorer definition of the point of onset of ionization at lower sensitivities. In any case, the effect of the finite spread in electron energies does not constitute the chief objection to the VC method. The chief objection is that the observed appearance potential, corresponding to the least energetic observable transition as described above, does not generally correspond to the energy difference between the lowest level of the ionized state in question and the lowest level of the ground state. If this limitation did not exist, there would probably be (apart from possible uncertainty in the electronic excitation) no discrepancy between VC electron impact data, on the one hand, and ionization potentials and dissociation energies obtained spectroscopically and thermochemically on the other hand.

General Significance of LE Appearance Potentials

For consideration of the significance of an appearance potential determined by the linear-extrapolation method it is convenient to make use of some simplifying assumptions which are not strictly correct; but they are probably not so extreme as to invalidate the analysis of the relationship of the various factors which influence an LE appearance potential. In the case of the ionization potential of an atom we will use in particular the basic assumption used recently by Vought[1] and earlier by others,[5] and indicated by the work of Stevenson and Hipple[3] to be not strictly correct near the critical potentials, namely, that the probability of ionization by electrons of energy $E_e > E_0$ is proportional to the difference $E_e - E_0$, where E_0 is the ionization potential. Let the distribution of electron energies be such that the number of electrons per second with energy E_e is given by $n_e = n_0 \cdot f_1(E_e - {}^0E_e)$, where n_0 is the number per second with the most probable energy ${}^0E_e$. Then the electrons in the energy interval dE_e near $E_e > E_0$ will, by causing ionization to the ${}^2P_{1\frac{1}{2}}{}^0$ state of ionized argon, say, produce an ion current:

$$I_1 = ck_1 \cdot (E_e - E_0) \cdot n_0 \cdot f_1(E_e - {}^0E_e) \cdot dE_e.$$

If we designate by α the energy difference between the ${}^2P_{\frac{1}{2}}{}^0$ and ${}^2P_{1\frac{1}{2}}{}^0$ states of the ionized atom, then electrons in the energy interval dE_e near $E_e > E_0 + \alpha$ will also, by causing ionization to the ${}^2P_{\frac{1}{2}}{}^0$ state of the ionized atom, produce an additional ion current:

$$I_2 = ck_2 \cdot (E_e - E_0 - \alpha) \cdot n_0 \cdot f_1(E_e - {}^0E_e) \cdot dE_e.$$

In these equations c is the instrumental sensitivity factor, which is really unimportant in this treatment, and k_1 and k_2 are the relative probabilities of the two processes under equally favorable energy conditions. When all the electrons have sufficient energy to cause ionization to either level, the total ion current at a particular ${}^0E_e$ will be given by:

$$I_t = c(k_1 + k_2) \int \left(E_e - E_0 - \frac{k_2 \alpha}{k_1 + k_2} \right) \cdot n_0 \cdot f_1(E_e - {}^0E_e) \cdot dE_e.$$

Writing $(E_e - {}^0E_e) + ({}^0E_e - E_0)$ for $(E_e - E_0)$, we obtain:

$$\frac{I_t}{c(k_1 + k_2)} = \int (E_e - {}^0E_e) \cdot n_0 \cdot f_1(E_e - {}^0E_e) \cdot dE_e + \int \left({}^0E_e - E_0 - \frac{k_2 \alpha}{k_1 + k_2} \right) \cdot n_0 \cdot f_1(E_e - {}^0E_e) \cdot dE_e.$$

If N is the total number of electrons per second, then the first integral becomes NA_e, where A_e is the average algebraic deviation of electron energy from the most probable electron energy; and the second integral is easily integrable, so that for sufficiently high electron energies we get the expression,

$$\frac{I_t}{c(k_1 + k_2)} = N \left({}^0E_e - E_0 - \frac{k_2 \alpha}{k_1 + k_2} + A_e \right),$$

which is linear in the most probable electron energy ${}^0E_e$. Linear extrapolation to $I_t = 0$ puts a condition on the most probable electron energy, namely,

$${}^0E_e{}^a = E_0{}^a + \frac{k_2 \alpha}{k_1 + k_2} - A_e, \qquad (1)$$

where the superscript, a, on the electron energy and ionization potential is an index to distinguish them from corresponding quantities in the subsequent discussion of an ion produced from a molecule. In practice the most probable electron energy is not measured directly, because of con-

[5] H. D. Smyth, Phys. Rev. **25**, 452 (1925).

tact potentials and other instrumental factors. What is measured is the nominal electron energy ${}^{a}E_e$, which differs from the most probable electron energy by an instrumental constant Q. Thus, the expression for the apparent LE appearance potential of an ion produced from a rare gas becomes

$$ {}^{a}E_e^{a} = E_0^{a} + \frac{k_2}{k_1+k_2} - A_e - Q. \qquad (1a) $$

Equation (1a) shows that the apparent LE ionization potential of a rare gas is influenced by the multiplicity of the lowest state of the ion as well as by asymmetry in the distribution of electron energies.

For the treatment of the LE appearance potential of an ion produced from a molecule we will speak in terms of a diatomic molecule for the sake of simplicity; but it will be clear that the results are more generally applicable. Let the ground level of the molecule be described by an expectation function such that $R_i = R_0 \cdot f_2(r_i - r_0)$ is the probability that the molecule has the internuclear distance r_i, where R_0 is the probability of occurrence of the most probable internuclear distance, r_0. Let the energy of a transition from the ground level of the molecule to the level of the ionic state which intersects the ionic state potential function at r_0 be E_0. Then the energies of transition to the other levels of the ionic state can be represented by $E_i = E_0 + f_3(r_i - r_0)$. We will assume the same type of dependence of transition probability on electron energy and the same electron energy distribution that we assumed in the case of ionization of a rare gas. In addition, we will assume, to a fairly good degree of approximation for unstable ionic states and to a sufficiently good degree of approximation for stable ionic states, that transitions occur from the ground level of the molecule only to the potential function of the ionic state itself, and that so far as the effect of the upper levels is concerned, all such possible transitions are equally probable under equally favorable energy conditions.

For unstable ionic states, the last assumption will be recognized as the same sort of approximation to the Franck-Condon principle as that introduced by Condon in one of his early papers.[6] It is equivalent to replacing the wave functions of the vibrational levels of the unstable ionic state by δ-functions of equal value at the turning points, so that the dependence of transition probability on the square of the overlap integral is reduced to dependence on the square of the wave function of the stable initial level. This approximation has been used frequently and has been shown by Coolidge, James, and Present[7] to yield results as satisfactory as strict application of the Franck-Condon principle in the case of some particular transitions in hydrogen. Of special interest is the fact that Hagstrum and Tate[2] have applied the approximation in successfully treating the contours of beams of ions produced in dissociation of diatomic molecules by electron impact, their function $(\psi_L)^2$ being equivalent to our R_i. It should be noted that the cited papers have dealt with cases in which one state was wholly unstable for the transitions considered. Application of the approximation to stable states, as done in the present paper, is not so well justified; but it is certainly good enough to lead to the proper qualitative conclusions.

Proceeding on the basis of the stated assumptions, it is clear that the ion current produced by electrons in the energy interval dE_e near $E_e > E_i$ as the result of their impact on molecules with internuclear distances in the interval dr_i near r_i will be given by

$$ dI = ck[E_e - E_0 - f_3(r_i - r_0)] \cdot n_0 \cdot f_1(E_e - {}^{0}E_e) \cdot R_0 \cdot f_2(r_i - r_0) \cdot dr_i \cdot dE_e. $$

Writing $(E_e - {}^{0}E_e) + ({}^{0}E_e - E_0)$ for $(E_e - E_0)$, we get for the total current of the particular ion at high electron energies:

$$ \frac{I_t}{ck} = \int\int (E_e - {}^{0}E_e) \cdot n_0 \cdot f_1(E_e - {}^{0}E_e) \cdot R_0 \cdot f_2(r_i - r_0) \cdot dr_i \cdot dE_e $$

$$ + \int\int ({}^{0}E_e - E_0) \cdot n_0 \cdot f_1(E_e - {}^{0}E_e) \cdot R_0 \cdot f_2(r_i - r_0) \cdot dr_i \cdot dE_e $$

$$ - \int\int f_3(r_i - r_0) \cdot n_0 \cdot f_1(E_e - {}^{0}E_e) \cdot R_0 \cdot f_2(r_i - r_0) \cdot dr_i \cdot dE_e. $$

[6] E. U. Condon, Phys. Rev. **32**, 858 (1928).
[7] A. S. Coolidge, H. M. James, and R. D. Present, J. Chem. Phys. **4**, 193 (1941).

Performing the integrations over the whole range of electron energies and over the range of internuclear distances appropriate to the ionization and dissociation process in question we get $I_i/Bck = N(A_e + {}^0E_e - E_0 - A_r)$, where N is the total number of electrons per second; A_e is the average algebraic deviation of the electron energy from the most probable electron energy, ${}^0E_e$; E_0 is the energy of the most probable transition; B is equal to $\int R_0 \cdot f_2(r_i - r_0) \cdot dr_i$ over the range of internuclear distances appropriate to the process; and A_r is the average algebraic deviation of energies of possible transitions (within this same range of internuclear distances) from the energy of the most probable transition, the transitions being weighted according to the expectation function of the ground level of the molecule.** The expression for the total ion current is thus linear in the most probable electron energy at high electron energies; and linear extrapolation to zero current requires that

$$ {}^0E_e^m = E_0^m + A_r - A_e, \qquad (2) $$

which, in turn, as in the case of Eq. (1a), leads to an expression for the apparent LE appearance potential:

$$ {}^aE_e^m = E_0^m + A_r - A_e - Q. \qquad (2a) $$

(The superscript, m, in the last two equations identifies the quantities as referring to the production of an ion from a molecule.) In practice, the difference between (2a) and the apparent LE ionization potential of a rare gas given in (1a) would be added to the known ionization potential of the rare gas to get the corrected appearance potential of the ion produced from the molecule. Performing these operations, we get for the corrected LE appearance potential of an ion produced from a molecule the value:

$$ \text{A.P.}(X^+) = E_0^m + A_r - \frac{k_2 \alpha}{k_1 + k_2}. \qquad (3) $$

It is to be noted, contrary to the implications of reference 1, that the asymmetry in the electron energy distribution does not influence the corrected appearance potential, inasmuch as the effects are equal for both the calibrating ion and the experimental ion. Thus the effects of curvature of the potential function of the ionized state in the Franck-Condon region and the effects of asymmetry of the expectation function of the initial state of the molecule remain in full measure, since both sets of effects are contained in A_r. In addition, the significance of LE appearance potentials is affected by the multiplicity of the lowest state of the ion of the rare gas used for calibration. To estimate the order of magnitude of the error introduced by neglect of the latter effect we can probably take the k's to be simply proportional to the quantum weights of the corresponding levels, since the last term in (3) is not extremely sensitive to the k's. Since α is 0.178 volts for argon,[8] the error would be about 0.06 volt when argon is used for calibration. If krypton, for which α is 0.67 volt,[8] is used for calibration, the error would be about 0.22 volt. It is desirable to note that consideration of the effect of multiplicity will probably not alter the conclusion of Stevenson and Hipple[3] with regard to the validity of the assumption of linear dependence of ionization probability on electron energy, since the discrepancy which they find between the true difference in the ionization potentials of neon and argon and the difference in the LE values will be reduced by only 0.03 volt.

Because of uncertainty in the second term in the expression (Eq. (3)) for an appearance potential determined by the linear-extrapolation method and of lack of precise knowledge concerning the details of the ionization process for electron energies near the critical potentials,[3] it can hardly be said that LE appearance potentials have a more definite physical significance than VC appearance potentials. There is a difference in significance, however, in that the LE appearance potentials correspond generally to a higher degree of vibrational excitation in the ionized state than do the VC appearance potentials. Thus the chief objection to the VC method is applicable in greater measure to the LE method. Another consideration favoring the VC method is that appearance potential curves for

** It will be observed that this treatment implies an essentially continuous distribution of levels for a stable ionic state as well as for an unstable one. Considerable simplification is gained in this way without impairing the qualitative conclusions.

[8] Bacher and Goudsmit, *Atomic Energy States* (McGraw-Hill Book Company, Inc., New York, 1932).

some ions do not become straight within a reasonable range of electron energies,[9] thus making application of the LE method impossible. Moreover, if an ion arises in two processes requiring energies so nearly the same that the appearance potential curve does not become straight before the onset of the more energetic process, the possibility of there being a second appearance potential may easily be missed if the LE method is applied, and the single value of appearance potential obtained by that method will not correspond to the value for either process, but rather to some intermediate value.

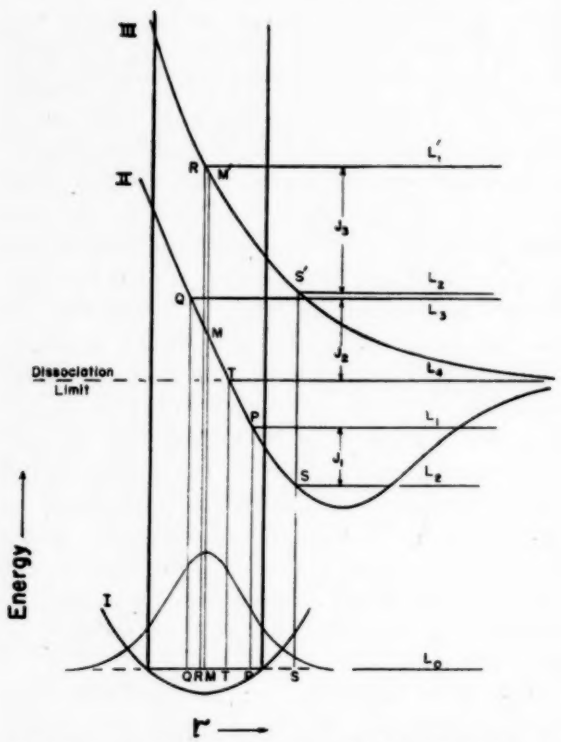


FIG. 1. Some hypothetical potential functions showing the approximate significance of corrected appearance potentials:

I	normal state of a hypothetical molecule.
II	stable ionic state of the molecule.
III	unstable ionic state of the molecule.
MM	energy of most probable transition between *I* and *II*.
MM'	energy of most probable transition between *I* and *III*.
SS	approximate *VC* appearance potential of molecular ions.
PP	the corresponding *LE* appearance potential $+k_2\alpha/k_1+k_2$.
TT	approximate *VC* appearance potential for fragment ions produced by transition to state *II*.
QQ	the corresponding *LE* appearance potential $+k_2\alpha/k_1+k_2$.
SS'	approximate *VC* appearance potential for fragment ions produced by transition to state *III*.
RR	the corresponding *LE* appearance potential $+k_2\alpha/k_1+k_2$.

[9] For example, see the curve for H^+ from methane; L. G. Smith, Phys. Rev. 51, 265 (1937).

Significance of the Difference between the LE and VC Appearance Potentials of an Ion

For a discussion of the possibility of deriving information concerning the configuration of the potential functions of the ionic states of a molecule from an investigation of the "feet" of the appearance-potential curves (i.e., the difference between the apparent LE and VC appearance potentials) we must first know the significance of the difference between the values of ${}^{a}E_e$ (or, alternatively, of ${}^{0}E_e$) which correspond to the LE and VC ionization potentials of the calibrating gas, since it is only by subtracting this difference from the corresponding difference for an ion produced from a molecule that we can make an approximate correction for the spread in electron energies and arrive at a quantity which we may hope to be related primarily to a potential function of an ionic state of the molecule. For the atomic calibrating ion the LE appearance potential corresponds to a particular condition for the most probable electron energy, which has already been stated in Eq. (1). The VC appearance potential, on the other hand, corresponds to a particular condition for ${}^{h}E_e$, the mean energy of the narrow energy band in which fall the first group of high energy electrons occurring in significant intensity in the electron energy distribution. This condition is that ${}^{h}E_e=E_0{}^{a}$, which, however, can be converted to a condition for the most probable electron energy, since ${}^{h}E_e$ is some definite energy, H, above ${}^{0}E_e$. The condition for ${}^{0}E_e$ corresponding to the VC appearance potential can thus be written ${}^{0}E_e{}^{a}=E_0{}^{a}-H$. The difference between the expression in Eq. (1) and this expression gives the energy interval between points corresponding to the VC and LE appearance potentials on the energy axis of the appearance-potential curve. This interval, which we will call the span, has then, for the calibrating ion, the value:

$$\text{span}_1=\frac{k_2\alpha}{k_1+k_2}+H-A_e. \qquad (4)$$

To help in interpreting the significance of the span for an ion produced from a molecule some hypothetical potential functions have been drawn in Fig. 1. Curve *I* is for the stable normal state of a hypothetical diatomic molecule, whose lowest vibrational level is L_0. The bold vertical lines represent the classical limits of vibration or the commonly accepted limits of the Franck-Condon region. The expectation function for the lowest level has been drawn schematically. Curves *II* and *III* represent stable and unstable ionic states, respectively. The potential functions *I*, *II*, and *III* are not intended to represent those of any real molecule, although, with certain features exaggerated, they are similar to potential functions for H_2 and H_2^+;[10] and the placement of levels in the figure is very rough, since that is all that is necessary for the subsequent discussion.

Let us consider transitions between the normal state and the stable ionic state represented by curve *II*. It is evident that such transitions can give rise to either a molecular ion or an atomic ion depending on the initial nuclear separation. We shall discuss first the molecular ion arising from transitions lying on the right of *TT*. The condition on ${}^0E_e$ corresponding to the LE appearance potential is given by Eq. (2) in which E_0^m is given by the ordinate *MM* in Fig. 1 and A_r is computed over the range of nuclear separation to the right of *TT*. The ordinate *PP* is drawn about equal to $A_r+E_0^m$, so that the LE appearance potential puts on the most probable electron energy the condition ${}^0E_e^m=PP-A_e$.*** The VC appearance potential puts a condition on the energy of the narrow band of high energy electrons, namely, that ${}^kE_e$ is approximately equal to an ordinate *SS* the position of which is not exactly defined;† and converting this to a condition on ${}^0E_e$ as was done in the case of the calibrating ion, we get for the span of the appearance-potential curve of the molecular ion,

$$\text{span}_2=J_1+H-A_e,$$

where $J_1=PP-SS$ in Fig. 1. Subtracting from this the span for the calibrating ion given in Eq. (4) we get

$$\text{span}_2-\text{span}_1=J_1-\frac{k_2\alpha}{k_1+k_2}. \qquad (5)$$

Thus, assuming that the second term on the right side of Eq. (5) can be taken into account, examination of the feet of the appearance-potential curves in this case will yield a value for J_1, which indicates nothing more than a minimum value for the energy range over which the potential function of curve *II* is stable. A better value for the minimum energy range of stability would be obtained by the difference between the VC appearance potential *TT* for the atomic ion and the VC appearance potential *SS* for the molecular ion, providing, of course, that it could be established that the two types of ions originated in transitions involving the same potential function. It should be observed that even this closer approximation would not give the energy difference between the dissociation limit and the minimum of curve *II*; for the level L_2, constituting the upper level of the least energetic of the observable transitions, is not necessarily the lowest vibrational level of the upper state.

If we consider the production of atomic ions by transitions to curve *II* corresponding to internuclear distances lying to the left of *TT*, we arrive at an expression for the difference between the span for the atomic ion and the span for the calibrating ion identical with Eq. (5), except that J_2 appears in place of J_1. Thus, assuming again that the effects peculiar to the calibrating ion can be taken into account, examination of the feet of the appearance-potential curves with which we are at present concerned will give us a value for J_2, which is indicative of a minimum value for the slope of curve *II* in the Franck-Condon region. The value would be extremely rough without prior knowledge of the range of internuclear distances between *QQ* and *TT*. In the present case J_2 would also indicate the total energy released in dissociation following the transition to the level L_3 corresponding to the LE appearance potential; but the foot of an appearance-potential curve cannot be given this unique interpretation without prior

[10] See curves *a*, *d*, and *f*, respectively, in Fig. 17 of Smyth's review article, Rev. Mod. Phys. **3**, 347 (1931).

*** From the definition of *PP* it is evident, by use of Eq. (3), that *PP* is also equal to the corrected LE appearance potential plus $k_2\alpha/k_1+k_2$.

† From the discussion of VC appearance potentials it is evident that the properly selected ordinate *SS* is equal to the corrected VC appearance potential.

knowledge of the relationships between the potential functions involved. For example, let us assume that the atomic ions produced by transitions to such a state as that represented by curve *III* are different from those just discussed, so that they can be distinguished in the mass spectrometer and their appearance-potential curve obtained. Examination of the foot of the curve as done in the other cases would yield a value for J_3. But this is only a fraction of the total energy released in dissociation following a transition to the level L_1' corresponding to the LE appearance potential; it is only an increment above that released in dissociation from the level L_2' corresponding to the VC appearance potential. It may, however, in distinction with the case just discussed, be rather closely indicative of the slope of the upper state potential function in the Franck-Condon region, providing that there is not considerable curvature in the function.

It is apparent that in the case in which the potential function of the upper state lies above the dissociation limit in the whole Franck-Condon region, as does curve *III*, the LE appearance potential of the fragment ion, properly corrected for the effect of the multiplicity of the lowest state of the calibrating ion, may give the transition energy in the center of the Franck-Condon region, uncertain to the extent of the effects of curvature of the upper state potential function and the effects of asymmetry in the expectation function of the ground level of the molecule, but independent of asymmetry in the electron energy distribution. However, for fragment ions produced by transitions to a state for which only part of the potential function lies above the dissociation limit, as illustrated by curve *II*, the LE appearance potential may correspond to a transition which is considerably removed from the center of the Franck-Condon region; and the range of curvature of the appearance-potential curve may not indicate even roughly the energy range of the upper state lying in the Franck-Condon region. Since it is generally unknown whether the upper state potential function is wholly or partly unstable in the Franck-Condon region, it seems clear, even within the limitations of the assumptions upon which the derivation of the significance of appearance potentials is based, that detailed examination of the foot of an appearance-potential curve is not likely to provide unambiguous information about the potential function of the ionic state involved. However, as the literature [2,10,11] shows, the magnitude of the appearance potentials themselves and other data obtainable by electron impact experiments can be helpful in testing proposals with regard to the nature of the potential functions of ionized states.

[11] For predictions concerning the results of electron impact studies of H_2, for example, see: Condon and Smyth, Proc. Nat. Acad. Sci. **14**, 871 (1928); Condon, Phys. Rev. **35**, 658A (1930); and for experiments verifying the predictions, see: Bleakney and Tate, Phys. Rev. **35**, 658A (1930); Bleakney, Phys. Rev. **35**, 1180 (1930); **40**, 496 (1932); Lozier, Phys. Rev. **36**, 1285 and 1417 (1930); **44**, 575 (1933).

Compton Scattering and the Emission of Low Frequency Photons

RES JOST*
Swiss Federal Institute of Technology, Zurich, Switzerland
(Received July 7, 1947)

According to the perturbation theory of the scattering of photons, the cross section for Compton scattering diverges, owing to the multiple scattering effects for low frequency photons. It is shown that by the application of the method of Block and Nordsieck this difficulty can be evaded, and that, to a first approximation, the cross section for Compton scattering is given by the Klein-Nishina formula.

1. INTRODUCTION

RECENTLY Eliezer[1] showed that by treating the multiple scattering of a photon colliding with a free electron one meets with a difficulty similar to that encountered in the theory of bremsstrahlung. This difficulty arises from the unjustified application of the perturbation method to low frequency photons. The cross section for double scattering becomes proportional to $\kappa^{-3}d^3\kappa d\Omega$, assuming that one quantum with wave number κ is scattered into the volume element $d^3\kappa$ of momentum space, while the other is scattered into the solid angle $d\Omega$. Integration over κ yields an infinitely large cross section for Compton scattering. We shall show that it is possible to escape this difficulty by performing a canonical transformation similar to that introduced by Bloch and Nordsieck[2] for the bremsstrahlung theory. To a first approximation the cross section for Compton scattering is then given by the Klein-Nishina formula.

2. CROSS SECTION FOR DOUBLE SCATTERING ACCORDING TO THE PERTURBATION THEORY

The process discussed here is the following: a primary photon $\mathbf{k}_0$ collides with a free electron with momentum $\mathbf{p}_0$ and energy T_0. In the final state, two photons $\mathbf{k}$ and $\boldsymbol{\kappa}$ are assumed to be present and the energy and momentum of the electron are T and $\mathbf{p}$, respectively.[3] Since we are only interested in the cross section for low frequencies of $\boldsymbol{\kappa}$, we can develop this quantity into a series of ascending powers of κ and retain only the lowest order term; i.e., the term in κ^{-3}, which is the only one that gives rise to difficulties. The matrix elements for the absorption of $\mathbf{k}_0$ and the emission of $\mathbf{k}$ and $\boldsymbol{\kappa}$ are, respectively,[4]

$$\left.\begin{aligned} H_{k_0} &= -e(2\pi/k_0)^{\frac{1}{2}}(u_1^*\alpha_0 u_2), \\ H_k &= -e(2\pi/k)^{\frac{1}{2}}(u_1^*\alpha u_2), \\ H_\kappa &= -e(2\pi/\kappa)^{\frac{1}{2}}(u_1^*\bar{\alpha} u_2), \end{aligned}\right\} \quad (1)$$

where $\alpha_0=(\boldsymbol{\alpha}\mathbf{e}_0)$, $\alpha=(\boldsymbol{\alpha}\mathbf{e})$, $\bar{\alpha}=(\boldsymbol{\alpha}\bar{\mathbf{e}})$, and $\mathbf{e}_0$, $\mathbf{e}$, $\bar{\mathbf{e}}$ are unit vectors in the direction of polarization of $\mathbf{k}_0$, $\mathbf{k}$, $\boldsymbol{\kappa}$, respectively; u_1 and u_2 are normalized spin functions. The matrix

$$H=\sum\frac{H_{AZ_1}H_{Z_1Z_2}H_{Z_2F}}{(E_A-Ez_1)(E_A-Ez_2)} \quad (2)$$

is responsible for the transition probability. The sum has to be extended over all permutations of the three processes as well as over both spin orientations and signs of the kinetic energy for the electron in the intermediate states. Developing Eq. (2) in a power series in κ, we need retain only the lowest order term proportional to κ^{-1}. From (1) it is then evident that the denominator in (2) has to be proportional to κ. This is only possible if the quantum $\boldsymbol{\kappa}$ is either emitted at the beginning or at the end of the sequence of the three processes, and then only if the intermediate state belongs to a positive energy of the electron. If, for example, $\boldsymbol{\kappa}$ is emitted at the beginning, we have the following states:

$$\begin{aligned} Z_1&: \quad [\mathbf{p}_1=\mathbf{p}_0-\boldsymbol{\kappa};\ \mathbf{k}_0,\ \boldsymbol{\kappa}], \\ Z_2'&: \quad [\mathbf{p}_2'=\mathbf{p}_0+\mathbf{k}_0-\boldsymbol{\kappa};\ \boldsymbol{\kappa}], \\ Z_2''&: \quad [\mathbf{p}_2''=\mathbf{p}_0-\mathbf{k}-\boldsymbol{\kappa};\ \mathbf{k}_0,\ \mathbf{k},\ \boldsymbol{\kappa}]. \end{aligned}$$

* A part of this investigation was elaborated in conjunction with Mr. Bengt Holmberg from Lund, during his stay in Zurich.

[1] C. J. Eliezer, Proc. Roy. Soc. **A187**, 210 (1946).

[2] F. Bloch and A. Nordsieck, Phys. Rev. **52**, 54 (1937).

[3] Natural units will always be used: $\hbar=c=1$.

[4] W. Heitler, Quantum Theory of Radiation (2nd ed.) (Oxford University Press, London, 1944) p. 96.

From the above we find

$$E-Ez_1=-(p_0,\kappa)/T_0+\cdots$$

where

$$(p_0,\kappa)=T_0\kappa-(\mathbf{p}_0\boldsymbol{\kappa})$$

Further we have

$$H_{Az_1}=-e(2\pi/\kappa)^{\frac{1}{2}}(u_0^*\bar{\alpha}u_1)$$

where u_1 belongs to positive kinetic energy and is an eigenfunction of $[\boldsymbol{\alpha}(\mathbf{p}_0-\boldsymbol{\kappa})+\beta m]$. By treating $(\boldsymbol{\alpha}\boldsymbol{\kappa})$ as a small perturbation, one gets for u_1 a development into a power series in κ. We are only interested in the unpertubated eigenfunction, which is u_0. Then $(\check{\mathbf{e}}\mathbf{p}_0)/T_0$ can be substituted for $\bar{\alpha}$. The leading term to (2) therefore becomes

$$e(2\pi/\kappa)^{\frac{1}{2}}[(\check{\mathbf{e}}\mathbf{p}_0)/(p_0,\kappa)]H_C \tag{3}$$

where H_C is the matrix element of the Compton effect. In the same way, for the case in which $\boldsymbol{\kappa}$ is emitted at the end, we obtain

$$-e(2\pi/\kappa)^{\frac{1}{2}}[(\check{\mathbf{e}}\mathbf{p})/(p,\kappa)]. \tag{4}$$

The quantities T, $\mathbf{p}$ are, respectively, the energy and momentum of the electron in the final state F. Combining (3) and (4) we get

$$H=e(2\pi/\kappa)^{\frac{1}{2}}\left\{\frac{(\check{\mathbf{e}}\mathbf{p}_0)}{(p_0,\kappa)}-\frac{(\check{\mathbf{e}}\mathbf{p})}{(p,\kappa)}\right\}H_C+\cdots. \tag{5}$$

The corresponding cross section is

$$d\Phi=(2\pi)^{-3}d^3\kappa w(\boldsymbol{\kappa},\check{\mathbf{e}})d\Phi_C+\cdots \tag{6}$$

with

$$w(\boldsymbol{\kappa},\check{\mathbf{e}})=2\pi\kappa^{-1}e^2\left\{\frac{(\check{\mathbf{e}}\mathbf{p}_0)}{(p_0,\kappa)}-\frac{(\check{\mathbf{e}}\mathbf{p})}{(p,\kappa)}\right\}^2. \tag{6'}$$

$d\Phi_C$ denotes the cross section of Klein and Nishina.[5] Summing over both directions of polarization $\check{\mathbf{e}}$ gives

$$d\Phi=(2\pi)^{-3}d^3\kappa w(\boldsymbol{\kappa})d\Phi_C \tag{7}$$

with

$$w(\boldsymbol{\kappa})=2\pi\kappa^{-1}e^2\left\{\frac{p_0^2\sin^2\varphi_0}{(p_0,\kappa)^2}+\frac{p^2\sin^2\varphi}{(p,\kappa)}-2\frac{([\mathbf{p}_0,\boldsymbol{\kappa}][\mathbf{p},\boldsymbol{\kappa}])}{\kappa^2(p_0,\kappa)(p,\kappa)}\right\}, \tag{7'}$$

φ_0 is the angle $(\boldsymbol{\kappa},\mathbf{p}_0)$; φ is the angle $(\boldsymbol{\kappa},\mathbf{p})$.

[5] $d\Phi_C$ is the relativistically invariant cross section as defined for example by C. Møller: Kgl. Danske Vidensk. Selsk. Mat.-fys. Meddel. **23**, no. 1 (1945), Section 2.

Equation (7) is, as it must be, invariant under Lorentz transformations since $d^3\kappa w(\boldsymbol{\kappa})$ may be written in a form which is evidently Lorentz invariant.

$$2\pi e^2\kappa^{-1}d^3\kappa\left\{2\frac{(p_0,p)}{(p_0,\kappa)(p,\kappa)}-\frac{m^2}{(p_0,\kappa)^2}-\frac{m^2}{(p,\kappa)^2}\right\} \tag{7''}$$

with

$$(p_0,p)=T_0T-(\mathbf{p}_0\mathbf{p}).$$

In this section we gave a new derivation of Eliezer's result. At the same time we found that the matrix element for the emission of $\boldsymbol{\kappa}$ may be split into two parts

$$\left[H_\kappa=-e(2\pi/\kappa)^{\frac{1}{2}}\left(u_1^*\frac{(\check{\mathbf{e}}\mathbf{p})}{T}u_2\right)+H_\kappa''\right],$$

in such a way that only the first part is responsible for the infra-red difficulties. The second part, H_κ'', may be treated along the lines of the usual perturbation theory. This splitting can be done in a more convenient way by starting with the Hamiltonian of radiation theory[6]

$$H=(\boldsymbol{\alpha}\mathbf{p})+\beta m-e(\boldsymbol{\alpha}\mathbf{A}(\mathbf{z}))+\sum_\nu N_\nu k_\nu \tag{8}$$

where $\mathbf{A}(\mathbf{x})$ is the transverse part of the vector potential, $\mathbf{z}$, $\mathbf{p}$ are the position and momentum vectors of the electron, respectively, N_ν is the number of photons with wave number $\mathbf{k}_\nu$ and polarization $\mathbf{e}_\nu$.

Expanding $\mathbf{A}(\mathbf{x})$ in a fourier series, we get:

$$\mathbf{A}(\mathbf{x})=\sum(2k_\nu)^{-\frac{1}{2}}q_\nu\mathbf{A}_\nu(\mathbf{x})+\text{conj.} \tag{9}$$

with

$$\mathbf{A}_\nu(\mathbf{x})=\mathbf{e}_\nu(2\pi)^{-\frac{1}{2}}\exp[i(\mathbf{k}_\nu\mathbf{x}_\nu)], \tag{10}$$

$$[q_\nu,q_{\nu'}^*]=\delta_{\nu\nu'}\quad\text{and}\quad q_\nu^*q_\nu=N_\nu. \tag{11}$$

Introducing

$$\mathbf{a}_\nu=e(2\pi/k_\nu)^{\frac{1}{2}}\mathbf{e}_\nu, \tag{12}$$

the Hamiltonian reads

$$H=(\boldsymbol{\alpha}\mathbf{p})+\beta m-\sum_\nu(\boldsymbol{\alpha}\mathbf{a}_\nu)q_\nu\exp[i(\mathbf{k}_\nu\mathbf{z})]+\text{conj.}+\sum_\nu N_\nu k_\nu. \tag{13}$$

Now we split the photons into high and low frequency ones. The high frequency (h.f.) photons are denoted by $\mathbf{k}$ and by an index l whereas the low frequency (l.f.) photons are

[6] W. Heitler, reference 4, p. 91.

denoted by κ and a running index s; the number of l.f. photons is n_s. To define the l.f. photons, we require that $\kappa<\omega\ll k_0$ in the center of gravity system (c.g.s.). A more precise definition of the l.f. quanta will be given in section 3. We now perform a canonical transformation involving only the l.f. quanta: $F'=S^{-1}FS$ with

$$S=e[i\sum_s n_s(\boldsymbol{\kappa}_s\mathbf{z})] \tag{14}$$

which gives

$$\mathbf{p}'=\mathbf{p}+\sum_s n_s\boldsymbol{\kappa}_s, \tag{15}$$

$$q_s'=q_s\exp[i(\boldsymbol{\kappa}_s\mathbf{z})]. \tag{16}$$

Omitting the primes the transformed Hamiltonian becomes:

$$H=(\boldsymbol{\alpha}\mathbf{p})+\beta m-\sum_s(\boldsymbol{\alpha}\mathbf{a}_s)(q_s+q_s^*)+\sum_s n_s(\kappa_s-(\boldsymbol{\alpha}\boldsymbol{\kappa}_s))-\sum_t(\boldsymbol{\alpha}\mathbf{a}_t)q_t\times\exp[i(\mathbf{k}_t\mathbf{z})]+\text{conj.}+\sum_t N_tk_t. \tag{17}$$

The $\boldsymbol{\alpha}$ in the terms involving the l.f. quanta are split into two parts[7]

$$\boldsymbol{\alpha}=(\mathbf{p}/T)+\boldsymbol{\alpha}_1 \quad \text{with} \quad T=+(m^2+p^2)^{\frac{1}{2}}, \tag{18}$$

$$H=(\boldsymbol{\alpha}\mathbf{p})+\beta m-\sum_s[(\mathbf{pa}_s)/T](q_s+q_s^*)+\sum_s n_s(p,\kappa_s)/T-\sum_t(\boldsymbol{\alpha}\mathbf{a}_t)q_s\times\exp[i(\mathbf{k}_t\mathbf{z})+\text{conj.}+\sum_t N_tk_t+H_1, \tag{19}$$

$$H_1=-\sum_s(\boldsymbol{\alpha}_1\mathbf{a}_s)(q_s^*+q_s)-\sum n_s(\boldsymbol{\alpha}_1\boldsymbol{\kappa}_s). \tag{20}$$

It is easily seen that already the terms of (19) without H_1 give rise to the cross section (7). H_1 does not lead to any difficulties if treated along the lines of the usual perturbation method. Therefore we shall omit H_1 in the following.

3. THE BLOCH-NORDSIECK TRANSFORMATION

It is well known that the breakdown of the perturbation method in the low frequency range is closely connected with the fact that the probability for the emission of a finite number of l.f. photons under the effect considered (e.g. bremsstrahlung or Compton effect) vanishes. The perturbation method should be reliable if restricted to photons with higher energy. In fact, this is not generally true as, for example, Dancoff[8] showed in the case of bremsstrahlung. But here the failure is connected with the still unsolved difficulties which have their origin in the high frequency region of the spectrum. It is known that neither Dirac's new method of field quantization[9] nor the formalism of Heitler-Wilson lead to a solution of this problem.[10]

In spite of this, we shall restrict the Bloch-Nordsieck transformation to l.f. photons and use the perturbation method for higher frequencies. In fact, it does not help to mix the more formal difficulties connected with l.f. photons with the fundamental problems connected with high energy photons.

The Bloch-Nordsieck transformation consists of the following canonical transformation $F'=S^{-1}FS$ with $S=e^U$

$$U=\sum_s[(\mathbf{pa}_s)/(p,\kappa_s)](q_s-q_s^*) \tag{21}$$

leading to

$$q_s'=q_s-(\mathbf{pa}_s)/(p,\kappa_s), \tag{22}$$

$$n_s'=n_s-\frac{(\mathbf{pa}_s)}{(p,\kappa_s)}(q_s+q_s^*)+\frac{(\mathbf{pa}_s)^2}{(p,\kappa_s)^2}. \tag{23}$$

The last term in (23) gives rise to a self-energy:

$$-\sum_s(\mathbf{pa}_s)^2/T(p,\kappa_s)=-\frac{e^2}{(2\pi)^2}\int\frac{d^3\kappa}{\kappa^2}\,\frac{v^2\sin^2\vartheta}{1-v\cos\vartheta} \tag{24}$$

where we have used the definition of $\mathbf{a}_s$ as given in (12) and summed over both directions of polarization; $v=p/T$, ϑ is the angle between $\boldsymbol{\kappa}$ and $\mathbf{p}$. In the c.g.s. we have to integrate over $\kappa<\omega\ll k_0$ and obtain:

$$-\frac{e^2}{2\pi}\left(2-\frac{1-v^2}{v}\ln\frac{1+v}{1-v}\right)\omega, \tag{25}$$

which is of the order ωe^2 and therefore may be neglected. Further, it is consistent to neglect the energy-momentum balance for the l.f. quanta; i.e., to omit the sum $\sum_s n_s(p,\kappa)/T$. We shall see, that the expectation value of this term is small compared with k_0 in the c.g.s.

---

[7] p/T is not exactly the even part of $\boldsymbol{\alpha}$, since T is always positive.

[8] S. M. Dancoff, Phys. Rev. **55**, 959 (1939).

[9] W. Pauli and J. M. Jauch, Phys. Rev. **65**, 255 (1944); W. Pauli, Helv. Phys. Acta, **19**, 234 (1946).

[10] H. A. Bethe and J. R. Oppenheimer, Phys. Rev. **70**, 451 (1946).

The reduced Hamiltonian reads:

$$H=(\alpha\mathbf{p})+\beta m+\sum_i N_i k_i+V, \tag{26}$$

where

$$V=-\sum_i(\alpha\mathbf{a}_i)q_i e^U \exp[i(\mathbf{k}_i\mathbf{z})]e^{-U}+\text{conj.} \tag{27}$$

The primes have been dropped, and U is the function (21).

Transforming to plane waves for the electron one finds (Pauli and Fierz[11] formulas (17)–(21a)):

$$(\cdots n_s\cdots,\mathbf{p}|V|\cdots n_s'\cdots,\mathbf{p}')$$
$$=V_0\prod_s(n_s,\mathbf{p}|K|n_s',\mathbf{p}') \tag{28}$$

where V_0 is the interaction matrix of the usual radiation theory referring only to h.f. quanta, whereas

$$(n,\mathbf{p}|K|n',\mathbf{p}')=e^{-w/2}(n!n'!)^{\frac{1}{2}}$$
$$\times\sum_{\mu=0}^{\infty}\frac{(i\sqrt{w})^{n'-n+2\mu}}{(n-\mu)!\mu!(n'-n+\mu)!}, \tag{29}$$
$$w(\kappa,\mathbf{e})=\left[\frac{(\mathbf{p}'\mathbf{a})}{(p',\kappa)}-\frac{(\mathbf{p}\mathbf{a})}{(p,\kappa)}\right]^2$$

is the same function as defined in (6′). In Eq. (29) $1/\sigma!$ is zero if $\sigma<0$.

4. THE CROSS SECTION FOR THE COMPTON SCATTERING

We start with the reduced Hamiltonian (26) and interpret $\mathbf{p}$ as the momentum of the electron and n_s as number of the free l.f. photons, whereas the photons eliminated in (23) are bound photons giving rise to the self-energy (24). Along the lines of the usual perturbation theory we get the cross section for the effect, that a quantum $\mathbf{k}$ is scattered into the solid angle $d\Omega$ and that beside this n_s l.f. photons are emitted. Using

$$\sum_{n'}(n,\mathbf{p}|K|n',\mathbf{p}')(n',\mathbf{p}'|K|n'',\mathbf{p}'')$$
$$=(n,\mathbf{p}|K|n'',\mathbf{p}'')$$
$$d\Phi(\cdots n_s\cdots)=\prod_s(w_s^{n_s}/n_s!)e^{-w_s}\cdot d\Phi_C. \tag{30}$$

The emission of the l.f. quanta obeys a Poisson distribution, that is, they are emitted independently. As a matter of fact, the numbers n_s are restricted by conservation of energy and momentum (which we have neglected) but this restriction is not serious because κ is very small. Further we see that there is always an infinite number of l.f. photons emitted since $\sum_s w_s$ diverges. Always assuming that the quantum $\mathbf{k}$ is scattered into $d\Omega$ we get as the expectation value for the energy of the l.f. quanta

$$\epsilon=\sum_s\kappa_s w_s \tag{31}$$

and for the momentum

$$\pi=\sum_s\kappa_s w_s. \tag{32}$$

Using the expression for w_s and then summing over both directions of polarization, (31) and (32) becomes

$$\epsilon=\frac{e^2}{4\pi^2}\int\frac{a^3\kappa}{\kappa}\kappa\left[2\frac{(p_0,p)}{(p_0,\kappa)(p,\kappa)}-\frac{m^2}{(p_0,\kappa)^2}-\frac{m^2}{(p,\kappa)^2}\right], \tag{33}$$

$$\pi=\frac{e^2}{4\pi^2}\int\frac{d^3\kappa}{\kappa}\kappa\left[2\frac{(p_0,p)}{(p_0,\kappa)(p,\kappa)}-\frac{m^2}{(p_0,\kappa)^2}-\frac{m^2}{(p,\kappa)^2}\right]. \tag{34}$$

(ϵ,π) is a four-vector since we integrate over a relativistically invariant domain: $\kappa<\omega$ in the c.g.s. Equation (33) reads in the c.g.s.:

$$\epsilon(\theta)=(2e^2/\pi)\omega\left\{\frac{1+2\alpha^2}{\alpha(1+\alpha^2)^{\frac{1}{2}}}\times\ln[\alpha+(1+\alpha^2)^{\frac{1}{2}}]-1\right\}, \tag{35}$$

with $\alpha=p/m\sin\frac{1}{2}\theta$, θ angle between $\mathbf{k}_0$ and $\mathbf{k}$. For $\alpha\ll1$

$$\epsilon=8e^2\cdot\omega\cdot\alpha^2/3\pi.$$

In particular, for the non-relativistic case

$$\epsilon=(8e/3\pi)^2\cdot\omega\cdot v^2\sin^2\tfrac{1}{2}\theta,\quad v=p/m \tag{36}$$

while for $\alpha\gg1$

$$\epsilon=(2e/\pi)^2\omega\{2\ln2\alpha-1\}. \tag{37}$$

The expression $(2/\pi)\{\cdots\}$ in (37) is of the order of unity up to energies of 10^4 m in the laboratory system. In this energy range we may

[11] W. Pauli and M. Fierz, Nuovo Cimento **15**, 1 (1938).

therefore conclude

$$\epsilon \ll \omega (\ll k_0). \tag{38}$$

From (34) it may be seen that π lies in the plane $(\mathbf{p}_0, \mathbf{k})$ and is symmetrical with respect to these two vectors. One easily finds:

$$T_0\epsilon - (\mathbf{p}_0\pi) = T \cdot \epsilon(\pi) \sin^2\tfrac{1}{2}\theta \tag{39}$$

and from this

$$|\pi(\theta)| = \frac{T}{p} \frac{|\epsilon(\theta) - \epsilon(\pi) \sin^2\tfrac{1}{2}\theta|}{\cos\tfrac{1}{2}\theta} \ll \omega. \tag{40}$$

Now we have to show that, after performing the Bloch-Nordsieck transformation, it is possible to calculate the double Compton effect. At the same time, we get a lower limit for ω. As already mentioned, a consistent calculation of the e^6-correction to the Klein-Nishina formula is impossible since this correction diverges. From the point of view of the perturbation theory (as well as from experimental data), it is reasonable to assume (as long as damping effects are excluded) that corrections to the Klein-Nishina formula (including corrections from multiple scattering) are of the relative order e^2. It will be proved that under this assumption a closer determination of ω does not affect the results of physical interest, admitting errors of the relative order e^2. To calculate the double Compton scattering, we start with the Hamiltonian (26). We write the cross section for the emission of a second quantum k' ($k' > \omega$ in the c.g.s.) into the element d^3k' in the form

$$d\Phi'(\mathbf{k}') = Q(\mathbf{k}') d^3\mathbf{k}' d\Phi \tag{41}$$

with $d\Phi$ from (30). Assuming $k' \ll k_0$ in the c.g.s. we may expand $Q(k')$ in a power series in k' and restrict ourselves to the lowest order term. As in Section 2 we get:

$$Q(\mathbf{k}') d^3k' = \frac{d^3k'}{(2\pi)^3} w(\mathbf{k}', \mathbf{e}') + \cdots \tag{42}$$

with $w(\mathbf{k}', \mathbf{e}')$ as in (6'), $\mathbf{e}'$ denoting the direction of polarization of $\mathbf{k}'$. The correction to the Klein-Nishina formula due to the double scattering is:

$$\int \sum Q(\mathbf{k}') d^3k',$$

with the summation sign $\sum$ denoting summation over both directions of polarization. The term (42) gives a contribution of the order of magnitude

$$\int \sum \frac{d^3k'}{(2\pi)^3} w(\mathbf{k}', \mathbf{e}) \cong (\epsilon/\omega) \ln(k_0/\omega) \simeq e^2 \ln(k_0/\omega),$$

with ϵ determined from Eq. (35).

It is reasonable to assume that the other terms in (44) give contributions of the order of magnitude e^2. We have, therefore, the condition

$$e^2 \ln(k_0/\omega) \ll 1. \tag{43}$$

The value of ω can be chosen so small that the use of development (42) may be justified. If (43) is fulfilled we get for the probability that under a Compton effect, in which a quantum $\mathbf{k}$ is scattered into $d\Omega$ and a second quantum $\mathbf{k}'$ ($k' \ll k_0$ in the c.g.s.) with polarization $\mathbf{e}'$ is emitted into the element d^3k':

$$w(\mathbf{k}', \mathbf{e}') \cdot d^3k'/(2\pi)^3.$$

The expectation value of energy momentum emitted into d^3k' reads

$$\begin{aligned} &k' w(\mathbf{k}', \mathbf{e}') d^3k'/(2\pi)^3, \\ &\mathbf{k}' w(\mathbf{k}', \mathbf{e}') d^3k'/(2\pi)^3, \end{aligned} \tag{44}$$

which is identical with Eqs. (31) and (32). Consequently the choice of ω is arbitrary provided that condition (43) is satisfied. This discussion can also be developed by using the function $S(\epsilon, \pi)$ introduced by Pauli and Fierz.[12] We are interested in the cross section for the process that a quantum $\mathbf{k}$ is scattered into the solid angle $d\Omega$, while the sum of energy momentum of the l.f. quanta lies between ϵ and $\epsilon + d\epsilon$, π and $\pi + d\pi$. This cross section can be written in the form[13]

$$d\Phi = S(\epsilon, \pi) d\epsilon d^3\pi d\Phi_C \tag{45}$$

where $d\Phi_C$ is the cross section of Klein and Nishina, while

$$S(\epsilon, \pi) d\epsilon d^3\pi = \sum w_s^{n_s} e^{-w_s}/n_s!. \tag{46}$$

[12] W. Pauli and M. Fierz, reference 11, p. 13.

[13] It has to be noted that Pauli and Fierz take into account the correction of the cross section due to the energy loss ϵ, whereas we neglect the influence of energy-momentum loss (ϵ, π). Compare (26) where we have neglected the term $\sum n_s T^{-1}(p_s, \kappa_s)$.

The sum has to be extended over all quanta which fulfill the inequalities:

$$\epsilon \leq \sum n_s \kappa_s \leq \epsilon + d\epsilon, \quad \pi \leq \sum n_s \kappa_s \leq \pi + d\pi.$$

It is easily seen, that[14]

$$S(\epsilon, \pi) = (2\pi)^{-4} \int d\chi_0 d^3\chi \times \exp[i(\chi_0 \epsilon - \chi\pi)] \exp[f(\chi_0, \chi)], \quad (47)$$

with

$$f(\chi_0, \chi) = \int \sum (2\pi)^{-3} d^3\kappa w(\kappa, \mathbf{e}) \times \{\exp[-i(\chi_0 \kappa - (\chi\kappa))] - 1\}. \quad (48)$$

The sum $\sum$ again denoting summation over both directions of polarization. If (χ_0, χ) transforms like a four vector, $f(\chi_0, \chi)$ is invariant, as is $S(\epsilon, \pi)$. If we are only interested in the energy loss in the c.g.s. we may define

$$S(\epsilon) = \int d^3\pi S(\epsilon, \pi) = (1/2\pi) \int d\chi e^{i\chi\epsilon} e^{f(\chi)} \quad (49)$$

with

$$f(\chi) = \int \sum [d^3\kappa/(2\pi)^3] w(\kappa, \mathbf{e})(e^{-i\chi\kappa} - 1). \quad (50)$$

The integration over the angles yields (compare Eqs. (33) and (35)):

$$f(\chi) = C \int_0^\omega d\kappa (e^{-i\chi\kappa} - 1)/\kappa \quad (51)$$

with

$$C = \frac{2e^2}{\pi} \left\{ \frac{1+2\alpha^2}{\alpha(1+\alpha^2)^{\frac{1}{2}}} \ln[\alpha + (1+\alpha^2)^{\frac{1}{2}}] - 1 \right\}. \quad (52)$$

The integral (51) can be evaluated by using tabulated functions, but it is easier to introduce another method of cutting off at the frequency (ω): introducing the function $e^{-\kappa/\omega}$ we get

$$f(\chi) = C \int_0^\infty \frac{d\kappa}{\kappa} e^{-\kappa/\omega} (e^{-i\chi\kappa} - 1) = \ln(1 + i\chi\omega)^{-C} \quad (53)$$

which yields for $S(\epsilon)$:

$$S(\epsilon) = \omega^{-1} (\epsilon/\omega)^{C-1} e^{-\epsilon/\omega} / \Gamma(C). \quad (54)^{15}$$

As $C \ll 1$ it is easily seen that $S(\epsilon)$ decreases very rapidly as ϵ increases.

To discuss the influence of the choice of ω, we go back to Eqs. (47) and (48). Again we have to take into account the double scattering.

Writing the cross section in the form:

$$d\bar{\Phi} = \bar{S}(\epsilon, \pi) d\epsilon d^3\pi d\Phi_C \quad (55)$$

where $d\bar{\Phi}$ is the cross section for which apart from the quantum $\mathbf{k}$ which is emitted into $d\Omega$, there is an energy-momentum loss $(\epsilon, \pi; \epsilon + d\epsilon, \pi + d\pi)$ arising from emission of l.f. quanta κ *or* from additional emission of a quantum k'. Under condition (43) restricting k' to an interval $\omega \leq k' \leq \omega_1 \ll k_0$ in the c.g.s., we get:

$$\bar{S}(\epsilon, \pi) = S(\epsilon, \pi) + \int \sum S(\epsilon - \eta, \pi - \zeta) \times w(\zeta, \mathbf{e}') \delta(\eta - \zeta) d\eta d^3\zeta. \quad (56)$$

Putting

$$\bar{S}(\epsilon, \pi) = (2\pi)^{-4} \int d\chi_0 d^3\chi \times \exp[i(\chi_0 \epsilon - (\chi\pi))] G(\chi_0, \chi) \quad (57)$$

one finds

$$G(\chi_0, \chi) = \exp[f(\chi_0, \chi)] \left\{ 1 + \int \sum (2\pi)^{-3} d^3k' \times w(\mathbf{k}', \mathbf{e}') \exp[-i(\chi_0 k' - (\chi\mathbf{k}'))] \right\} \quad (58)$$

which, to a good approximation, is equal to

$$\exp[f(\chi_0, \chi) + f'(\chi_0, \chi)]$$

with

$$f'(\chi_0, \chi) = \int \sum (2\pi)^{-3} d^3k' w(\mathbf{k}', \mathbf{e}') \times \{\exp[-i(\chi_0 k' - (\chi\mathbf{k}'))]\}, \quad (59)$$

where $\omega \leq k' \leq \omega_1$ in the c.g.s.

Therefore we get the same result as if we had extended the domain of the l.f. quanta to $\kappa \leq \omega_1$, in the c.g.s.

The author wishes to express his sincere thanks to Professor W. Pauli, who suggested this problem, for many valuable discussions.

[14] Compare W. Pauli and M. Fierz, reference 11, Eq. (32).

[15] W. Pauli and M. Fierz, reference 11, p. 14, Eq. (36).

Induction Effects in Terrestrial Magnetism

Part III. Electric Modes

WALTER M. ELSASSER*
RCA Laboratories Division, Princeton, New Jersey
(Received March 27, 1947)

It can be shown that the currents in the earth's core which give rise to the externally observable magnetic field do not form a complete set of solutions of the field equations. There exists a second set of solutions composed of the modes of the electric type which produce a magnetic field inside the metallic sphere, but appear at the outside only through an electric field too weak to be measured. For reasons of symmetry the most important terms among the electric modes are the quadrupoles. The theory of inductive coupling by fluid motion, developed previously, is here applied to the interaction of the magnetic and electric modes. The system again is non-conservative, and work is done on the field by the fluid, or *vice versa*. It is shown that the interaction between the magnetic dipole and electric quadrupole modes constitutes a basic amplifier mechanism which amplifies the quadrupole mode until the magneto-mechanical forces exerted by the field upon the fluid begin to slow down the motion, thus prohibiting further increase of the field. This internal quadrupole field is likely to be much larger than the ordinary magnetic dipole field. Further analysis leads one to interpret the couplings between the magnetic and electric modes as a feed-back amplifier whereby the field can be maintained through the power delivered to it by the fluid motion. A survey of possible sources of power for this process indicates that the power for the maintenance of the field is provided from the rotational energy lost by the earth as it is slowed down through the action of the lunar tide.

INTRODUCTION

THE analysis of Part I[1] and Part II[2] has led to an interpretation of the geomagnetic secular variation in terms of interactions between fluid motions in the earth's metallic core and electric currents in the core that are the sources of the magnetic field. This analysis suffers from the shortcoming that the current modes which give rise to a magnetic field outside the metallic sphere do not represent a complete set of solutions of the electromagnetic field equations. There exists a second set of solutions, representing modes of the electric type, whose magnetic field is confined to the interior of the conducting sphere. In the preceding parts these modes have been disregarded on the assumption that they cannot be excited. It has been found, however, that in the theory of inductive coupling by fluid motion there appear definite couplings between the two types of modes and that, therefore, the electric modes are an integral part of the field as described by this theory. It will appear in the course of this paper that from this viewpoint inductive coupling between the magnetic and electric modes is by far the most important feature of the earth's magnetic field.

FREE ELECTRIC MODES

It is a known fact[3] that a sphere of macroscopic dimensions and of a conductivity of the metallic order has four distinct sets of electromagnetic modes of free oscillations. There are two high frequency modes, oscillations that correspond to wave-lengths comparable to the radius of the sphere. These modes are not of interest to us here. There are two low frequency modes which, in the case of a metallic sphere of macroscopic size, are completely aperiodic. The free aperiodic modes of the magnetic type have been extensively dealt with in Part I. The present account of the aperiodic modes of the electric type can therefore be brief; the procedure and symbols follow those of Part I. We first give a treatment in which the displacement current is omitted from the outset as has been done in the case of the magnetic modes. This method is extremely simple, but leads to an apparent inconsistency. The essential correctness of the result will then be demonstrated by deriving them from the more general solutions that include the displacement current.

* Now at Randal Morgan Laboratory of Physics, University of Pennsylvania, Philadelphia, Pennsylvania.
[1] W. M. Elsasser, Phys. Rev. **69**, 106 (1946), designated as Part I in the text.
[2] W. M. Elsasser, Phys. Rev. **70**, 202 (1946), designated as Part II in the text.
[3] P. Debye, Ann. d. Physik **30**, 57 (1909).

The relation of the magnetic and electric modes may be described by saying that the symmetry types of the field vectors are interchanged. If again **T** and **S** designate the toroidal and poloidal vector fields defined in Part I, we have

magnetic modes: $\mathbf{B}\sim\mathbf{S}$, $\mathbf{E}\sim\mathbf{T}$
electric modes: $\mathbf{B}\sim\mathbf{T}$, $\mathbf{E}\sim\mathbf{S}$.

In analogy to (I, 26) we set now for the vector potential of any individual free mode, inside the sphere

$$\mathbf{A}^{(i)}=c\mathbf{S}e^{-\Lambda t}, \tag{1}$$

where all the pertinent relations of Part I, namely, (I, 8), (I, 11), (I, 24), (I, 27) remain unchanged. We quote the last-named

$$\Lambda=k^2/\mu\sigma. \tag{2}$$

We then find readily for the field vectors in the interior:

$$\mathbf{B}^{(i)}=cRk^2\mathbf{T}e^{-\Lambda t}, \quad \mathbf{E}^{(i)}=c\Lambda\mathbf{S}e^{-\Lambda t}. \tag{3}$$

The field equations in the space external to the sphere are fulfilled if we set

$$\mathbf{B}^{(e)}=0, \quad \mathbf{E}^{(e)}=C\Lambda R^n\nabla(r^{-n-1}Y_n^m)e^{-\Lambda t}. \tag{4}$$

The boundary conditions are the same as in the case of the magnetic modes. They readily yield the relation

$$J_{n+\frac{1}{2}}(k_{ns}R)=0, \tag{5}$$

which is to be compared to the condition (I, 16) for the magnetic modes.

Again we wish to normalize these modes in analogy to the normalization of the magnetic modes which was given by (I, 21). Hence, we require

$$\int\mathbf{S}\cdot\mathbf{S}^*dV=1,$$

the integral extending over the interior of the sphere. The calculation shows[4] that the normalization of the spherical harmonics (I, 22) may be left the same, but the normalized radial functions are now, in place of (I, 23),

$$Z_n(k_sr)=2^{\frac{1}{2}}(2n+1)^{\frac{1}{2}}R^{-2}[(n+1)J_{n-\frac{1}{2}}(k_sR)+nJ_{n+3/2}(k_sR)]^{-\frac{1}{2}}r^{-\frac{1}{2}}J_{n+\frac{1}{2}}(k_sr). \tag{7}$$

[4] By (I, 18); see also J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), Section 7.13.

Since the tangential component of **E** is continuous, we can readily estimate the magnitude of the external electric multipole field given by the second Eq. (4). The electrostatic potential, U, of this field is of the order

$$U\approx RE\approx RC\Lambda\approx Rc\Lambda\approx B/\mu\sigma.$$

Taking the magnetic field to be 1 gauss $=10^{-4}$ m.k.s. units, we find U to be of the order of 10^{-4} volt. Such a field cannot, of course, be detected by electrostatic measurements at the surface of the earth.

The preceding derivation of the electric modes suffers from a flaw that might be stated as follows. Since the radial component of **E** vanishes when the boundary is approached from the inside, there must be a charge on the surface of the sphere (at least for any finite assumed value of the dielectric constant on the inside). As this charge must decay exponentially in time, there should be a current to or from the surface, in contradiction to the boundary condition for **E** just stated. Closer scrutiny indicates that this current is small of the second order in the neglected displacement-current term. To show this, we shall derive the aperiodic modes from the more general case of the oscillatory modes. The characteristic equation for the electric modes is[5] (assuming μ constant throughout)

$$\psi_n'(k_iR)/\psi_n(k_iR)=k_i\zeta_n'(k_eR)/k_e\zeta_n(k_eR), \tag{8}$$

where

$$\psi_n(x)=x^{\frac{1}{2}}J_{n+\frac{1}{2}}(x), \quad \zeta_n(x)=x^{\frac{1}{2}}H_{n+\frac{1}{2}}^{(1)}(x).$$

$H^{(1)}$ being the Hankel function of the first kind. The propagation constants, k_i and k_e, have the usual meaning and refer to the internal and external space, respectively. Both are given functions of the frequency. If now the right-hand side of (8) is reduced to the first term of a power series in k_eR, one obtains

$$\psi_n(k_iR)=-(Rk_e^2/nk_i)\psi_n'(k_iR)$$

which for small values of k_e goes over into $\psi_n(k_iR)=0$, identical with (5). If the roots of this equation are designated $(k_iR)_0$ we have in the next approximation

$$k_iR=(k_iR)_0(1-k_e^2/nk_i^2)_0.$$

[5] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), Section 9.22.

The second term in the parenthesis now gives rise to a small current normal to the boundary.

INDUCTIVE COUPLINGS

For the sake of completeness we start again from the field equations

$$\nabla \times \mathbf{B} = \mu\sigma\mathbf{E} + \mu\sigma\mathbf{v} \times \mathbf{B}, \tag{9}$$

$$\nabla \times \mathbf{E} + \partial\mathbf{B}/\partial t = 0, \quad \nabla \cdot \mathbf{B} = 0. \tag{10}$$

We shall moreover require the energy integral. Apply to (9) and (10) the conventional procedure of deriving Poynting's theorem and neglect the surface integral of the Poynting vector when the surface over which the latter is taken tends to infinity. After some simple transformations we get

$$\tfrac{1}{2}(\partial/\partial t)\int \mathbf{B}^2 dV = \int \nabla \cdot (\mathbf{B} \times (\nabla \times \mathbf{B}))dV - (\mu\sigma)^{-1}\int (\nabla \times \mathbf{B})^2 dV, \tag{11}$$

where the integral on the left extends over all space, those on the right over the sphere only. The last term is proportional to the square of the *total* current as given by the right-hand side of (9), and represents the rate at which heat is generated. The first term on the right-hand side of (11) represents the change in the field energy produced by the motional induction. As this term does not, in general, vanish, the system is non-conservative; work is being done by the fluid on the field, or *vice versa*.

As we are primarily interested in the motional induction we shall assume $\mathbf{v}$ and σ so large that the free decay terms (and diffusion effects treated in Part II) become negligible. The transition to this limit does not involve any mathematical or physical difficulties. Equation (9) can then be written in the simple form

$$\partial\mathbf{A}/\partial t = \mathbf{v} \times \mathbf{B} = \mathbf{v} \times (\nabla \times \mathbf{A}). \tag{12}$$

Following the developments of earlier parts, put

$$\mathbf{A} = \sum_\gamma c_\gamma \mathbf{A}_\gamma, \quad \mathbf{v} = \sum_\alpha v_\alpha \mathbf{V}_\alpha \tag{13}$$

where the vectors with indices are of the orthogonal set, and where, moreover, the summation extends over both types of vectors, $\mathbf{T}$ and $\mathbf{S}$. From now on, we shall use *real* vectors in place of the complex vectors of the former papers; the coefficients in (13) are then real, and the normalizations (6) and (I, 21) are replaced by real conditions (increasing the normalization factor (I, 22) by $\sqrt{2}$, if $m \neq 0$).

Equation (12) becomes now

$$\dot{c}_\gamma = \sum_{\alpha\beta} v_\alpha c_\beta [\mathbf{V}_\alpha \cdot \mathbf{B}_\beta \times \mathbf{A}_\gamma], \tag{14}$$

which agrees with the former relations (I, 34) or (II, 10) except for the fact that the summation is now understood to extend over both types of modes. The expression (11) for the magnetic energy can be put in an analogous form. For any individual mode we have

$$\nabla \times \mathbf{B}_\gamma = \nabla \times \nabla \times \mathbf{A}_\gamma = k_\gamma^2 \mathbf{A}_\gamma$$

and hence (11) may be written, on neglecting the last integral,

$$dE_{\text{mag}}/dt = \sum_{\alpha\beta\gamma} v_\alpha c_\beta c_\gamma k_\gamma^2 [\mathbf{V}_\alpha \cdot \mathbf{B}_\beta \times \mathbf{A}_\gamma]. \tag{15}$$

On substituting from (14) and integrating, we find

$$E_{\text{mag}} = \tfrac{1}{2}\sum_\gamma k_\gamma^2 c_\gamma^2. \tag{16}$$

Again, let $\mathbf{F}$ be the density of the magneto-mechanical forces exerted by the field upon the fluid. Let $\mathbf{F}_\alpha$ be an orthogonal vector set; then

$$\mathbf{F} = \mathbf{J} \times \mathbf{B} = -\sigma(\partial\mathbf{A}/\partial t) \times \mathbf{B} = -\mu^{-1}\sum_{\alpha\beta\gamma} c_\beta c_\gamma k_\gamma^2 [\mathbf{F}_\alpha \cdot \mathbf{B}_\beta \times \mathbf{A}_\gamma]\mathbf{F}_\alpha \tag{17}$$

the last equality by (1) and (2).

We now proceed to classify all the matrix elements appearing in (14) according to the types of vectors, $\mathbf{T}$ or $\mathbf{S}$, involved. In the "primary" mode (index β) the magnetic field itself enters, in the "secondary" mode (index γ), the vector potential appears instead. We may distinguish the following types:

Interactions among magnetic modes:

$$[\mathbf{T}_\alpha \cdot \mathbf{S}_\beta \times \mathbf{T}_\gamma], \quad [\mathbf{S}_\alpha \cdot \mathbf{S}_\beta \times \mathbf{T}_\gamma]. \tag{18a, b}$$

Interactions among electric modes:

$$[\mathbf{T}_\alpha \cdot \mathbf{T}_\beta \times \mathbf{S}_\gamma], \quad [\mathbf{S}_\alpha \cdot \mathbf{T}_\beta \times \mathbf{S}_\gamma]. \tag{19a, b}$$

Magnetic mode primary, electric mode secondary:

$$[\mathbf{T}_\alpha \cdot \mathbf{S}_\beta \times \mathbf{S}_\gamma], \quad [\mathbf{S}_\alpha \cdot \mathbf{S}_\beta \times \mathbf{S}_\gamma]. \tag{20a, b}$$

Electric mode primary, magnetic mode secondary:

$$[\mathbf{S}_\alpha \cdot \mathbf{T}_\beta \times \mathbf{T}_\gamma]. \qquad (21)$$

In the case (21) there is only one type of interaction, since all matrix elements containing three toroidal vectors vanish. All the preceding matrix elements differ from the two types (18) only by permutations, with the exception of (20b). The types (18) have been described in the preceding parts and a table of all the elements containing only dipole and quadrupole vectors has been given.[6] The type (20b) is new; a table of the lowest elements will be found in the appendix. This type appears to be of minor physical significance, and we shall not encounter it later on. From the tables, or the analytical expressions given in the appendix, the values of *all* matrix elements (18)–(21) can be obtained either directly or by suitable permutations of the vectors or their indices.

Next, we try to simplify Eq. (14). We begin by imposing two restrictions on the fluid motion, namely, assuming rotational symmetry about the earth's axis and symmetry about the equatorial plane. The first condition restricts us to zonal vectors ($m=0$); this will later be somewhat relaxed. The second condition requires that the poloidal flow components, $\mathbf{S}_n$, must have even n, the toroidal flow components, $\mathbf{T}_n$, must have odd n. This is readily verified from the expressions (I, 13) and (I, 14) for these vectors. In Part I we derived certain selection rules for the matrix elements; from the formulas of the appendix these can now be generalized to apply to all the elements (18)–(21): For the elements (18b), (19b), and (20a), containing two vectors **S** and one vector **T**, the sum of the three indices n must be even; for the remaining elements containing either one vector **S** and two vectors **T**, or else three vectors **S**, the sum of the three indices n must be odd. One finds now that under the symmetry assumptions made for the fluid motion, the entire set of electromagnetic modes is *reducible* into two sets that are not coupled to each other. One of these, the symmetrical set, has electric currents or vector potentials which have the same symmetry as the fluid motion, i.e., for $\mathbf{A}=\mathbf{S}_n$ the index n must be even, for $\mathbf{A}=\mathbf{T}_n$ it must be odd. The other, antisymmetrical set, is complementary to the symmetrical set; the electric currents or vector potentials are antisymmetrical about the equatorial plane. In order to prove this we need only insert the corresponding vectors into the matrix elements (18)–(21); it then appears that the couplings between the two sets vanish by virtue of the selection rules just enunciated. Since these selection rules do not depend on the magnitude of the tesseral index, m, we can at once generalize the result so that it applies to tesseral harmonics as well: Let the symmetrical set of vectors be composed of the $\mathbf{S}_n^m$ where n is even and of the $\mathbf{T}_n^m$ where n is odd; the antisymmetrical set is complementary to this. Then, when the fluid motion can be described as a linear aggregate of vectors of the symmetrical set, the electromagnetic field is reducible; modes whose electric currents (or vector potentials) are of the symmetrical set are not coupled to modes whose electric current are of the antisymmetrical set.

It is now furthermore found from (17) that the magneto-mechanical forces engendered by any one set of modes alone are vectors of the symmetrical set; the forces engendered by the interaction of a symmetrical and an antisymmetrical mode are vectors of the antisymmetrical set. It is clear that the primary mechanical forces acting upon the fluid, at least so long as they have rotational symmetry, must belong to the symmetrical set. This is evident for any forces caused by pressure differences between the polar and equatorial regions. The Coriolis force is of the form $\boldsymbol{\omega} \times \mathbf{v}$, where $\boldsymbol{\omega}$ is a vector of type $\mathbf{S}_1$ derived by (I, 14) from a generating function $\psi = (r/2R)\omega \cdot \cos\vartheta$. The development of $\boldsymbol{\omega} \times \mathbf{v}$ in terms of the fundamental vectors is then seen to yield only vector components of the symmetrical set. Thus we see that the forces, velocities and electric currents pertaining to the symmetrical set form a self-consistent system. This is not the case when the electric current modes are of the antisymmetrical set, since the resultant magneto-mechanical forces are symmetrical. Observation shows that the magnetic dipole modes, which are symmetrical vectors, are large, while the quadrupoles, which are antisymmetrical vectors, are very small. It will henceforth be assumed that the symmetrical set of electric

[6] Reference 2, p. 207.

current modes is the one that preponderates in the earth. This leads at once to the conclusion that for the modes of the electric type the quadrupoles must be large and the dipoles and octupoles small.

On relinquishing our original requirement of rotational symmetry for the field vectors, it will be appropriate to retain the assumption that the deviations from this symmetry are small. The simplest such case is that of an axis of rotation slightly inclined relative to the earth's axis. Let ϵ, ψ be the polar angles of the new axis in the old system and let ϑ' be measured from the new axis. From the addition theorem of the spherical harmonics we get, up to terms quadratic in ϵ,

$$P_n(\cos\vartheta') = (1-\tfrac{1}{2}\epsilon^2)P_n(\cos\vartheta) + \epsilon P_n^1(\cos\vartheta)\cos(\varphi-\psi) + (\epsilon^2/4)P_n^2(\cos\vartheta)\cos 2(\varphi-\psi)+\cdots, \quad (22)$$

where the P's are the ordinary (unnormalized) Legendre polynomials. From this formula the relations for the corresponding vectors containing normalized spherical harmonics are readily derived. In what follows, we shall confine ourselves to first-order terms in ϵ, that is to tesseral harmonics with $m=1$.

In order to achieve further simplification of Eq. (14) we shall now limit ourselves to the dipoles of the magnetic modes and the quadrupoles of the electric modes. We have, therefore, the following six vectors for the electric currents or vector potentials of the large modes

$$\mathbf{T}_1, \mathbf{T}_1^c, \mathbf{T}_1^s; \quad \mathbf{S}_2, \mathbf{S}_2^c, \mathbf{S}_2^s, \quad (23)$$

where the upper indices c and s represent the functions $\cos\vartheta$ and $\sin\vartheta$, respectively. Equation (14) reduces now to six sets (where the members of each set are distinguished only by the index of the radial eigenfunction). In place of the ordinary differential Eq. (14) we could form a set of *partial* differential equations for the two independent variables r and t; under the restriction (23) these would reduce to six simultaneous partial differential equations (or four simultaneous complex equations).

The assumption of rapid convergence of the spherical harmonic series is, of course, justified by the fact that the higher harmonic components are quickly damped out by free decay. There is no similar restriction on the components of the fluid motion. We can, however, limit ourselves to the vectors

$$\mathbf{T}_1, \mathbf{T}_1^c, \mathbf{T}_1^s; \quad \mathbf{S}_2, \mathbf{S}_2^c, \mathbf{S}_2^s; \quad \mathbf{T}_3, \mathbf{T}_3^c, \mathbf{T}_3^s \quad (24)$$

as it follows from the selection rules of Part I that, under the restriction (23) for the electric current vectors, only the velocity vectors (24) give rise to non-vanishing matrix elements.

AMPLIFICATION

We shall now study the individual interaction terms that involve vectors of the types (23) for the electric currents and (24) for the velocities. First, consider matrix elements that have full rotational symmetry. There are only four types of these

$$[\mathbf{S}_2 \cdot \mathbf{S}_1 \times \mathbf{T}_1], \quad [\mathbf{S}_2 \cdot \mathbf{T}_2 \times \mathbf{S}_2], \quad (25a, b)$$

$$[\mathbf{T}_1 \cdot \mathbf{S}_1 \times \mathbf{S}_2], \quad [\mathbf{T}_3 \cdot \mathbf{S}_1 \times \mathbf{S}_2], \quad (26a, b)$$

each element standing for an infinite sequence distinguished by the indices of the radial eigenfunctions. The elements (25a) represent interactions among the magnetic dipole components which have been discussed at some length in Part II. The elements (25b) represent interactions among the electric quadrupole components of a closely similar character. Our attention will be centered on the elements (26) which represent an induction from the magnetic dipole as primary to the electric quadrupole as secondary. These four types of elements exhaust the couplings of rotational symmetry among the rotationally symmetrical magnetic dipole and electric quadrupole modes. *There is, therefore, no reverse to the interactions (26).* This statement holds more generally in the sense that there are interactions of rotational symmetry between magnetic modes as primaries and electric modes as secondaries, but no interactions of rotational symmetry in the reverse direction. The proof can readily be deduced from the expressions of the matrix elements in the appendix.

For simplicity, let now the series of matrix elements resulting from the different radial eigenfunctions be represented symbolically by a single element; then Eq. (14) reduce to

$$dc_2/dt = [\mathbf{T}_1 \cdot \mathbf{S}_1 \times \mathbf{S}_2]v_1c_1 + [\mathbf{T}_3 \cdot \mathbf{S}_1 \times \mathbf{S}_2]v_3c_1. \quad (27)$$

If we assume that $\mathbf{v}_1$ and $\mathbf{v}_3$ are constants, and that the primary field is constant, the amplitude of the electric quadrupole increases linearly with time. The process (27) constitutes a basic mechanism of amplification. In somewhat more physical terms, this process might be described as follows: The primary dipole field, $\mathbf{S}_1$, is in the meridional plane, the velocity, $\mathbf{T}_1$ or $\mathbf{T}_3$, is perpendicular to this plane. The induced secondary current, being the vector product of the two, must again be in the meridional plane. The lines of magnetic force of the quadrupole thus generated are circles about the earth's axis. Hence, the secondary magnetic field is parallel to the velocity and there is no interaction between them which would produce a tertiary field. This is the reason why, in the presence of only the fields and velocities appearing in (27), Eq. (14) is *reducible* to the latter; the process of amplification of the quadrupole mode does not necessitate the simultaneous generation of other modes.

Now Eq. (27) is symbolical in the sense that in reality there is an infinite sequence of matrix elements corresponding to the different radial eigenfunctions of the primary as well as the secondary mode. But it is seen from the preceding geometrical consideration that on any circle $r=\text{const.}$, $\vartheta=\text{const.}$ the field of the secondary mode increases linearly with time, although the rate of increase will, in general, be a function of r and ϑ. This linear increase must continue until the magneto-mechanical forces exerted by the field upon the fluid become so large that they decelerate the fluid motion, thereby prohibiting further amplification. Equation (15) gives the amount of power delivered by the fluid motion to the field in the process of amplification. By (17) there corresponds to each matrix element a force component whose direction is everywhere opposite to the corresponding velocity component. In Part II the critical field strength for which the magneto-mechanical forces become equal in magnitude to the purely mechanical forces has been estimated. In view of (17), the last equation of Part II (preceding the appendix) may be modified to read

$$(B_1B_2)^{\frac{1}{2}}\sim(2\omega\rho/\sigma)^{\frac{1}{2}}=12 \text{ gauss}, \qquad (28)$$

where B_1 and B_2 designate the primary and secondary field, respectively. For larger fields than given by (28) the magneto-mechanical reactions preponderate, for smaller fields, the purely mechanical driving forces. B_1 in the upper layers of the core is about 2–3 gauss which would make B_2 about 50–60 gauss. The force exerted upon the velocity components $\mathbf{T}_1$ and $\mathbf{T}_3$ is, however, not the largest force exerted by the electromagnetic field. It results from the interaction of the secondary current with the primary field; a larger force results from the interaction of the secondary current with the secondary field. The latter force component is of the type $\mathbf{S}_2$. Whether or not this force influences the fluid motions in such a way that it eventually tends to decrease the motion of the types $\mathbf{T}_1$ and $\mathbf{T}_3$ indirectly, could only be decided by a more extensive analysis. In order to fix our ideas we shall assume a specific figure for the mean value of the electric quadrupole field in the core, namely, 30 gauss. This is presumably fairly close to a lower limit for reasons which will appear later.

The effects just mentioned must have a rather profound influence upon the hydrodynamics of the motions in the core. Since the magneto-mechanical reactions are large in certain directions, the problem is different from one of ordinary hydrodynamics. Instead, we are dealing with a system where the coupling between the fluid motion and the magnetic field is extremely close and their mutual interactions cannot be neglected. It is interesting to note that, in the case of sunspots, Alfvén[7] has arrived at closely analogous conclusions. He finds that in sunspots the magneto-mechanical forces must be greatly in excess of all mechanical forces; hence the motion must be controlled by the former. The ideas of Gurevich and Lebedinsky[8] on sunspots are in substantial agreement with this view.

FEEDBACK

So far, we have dealt with interactions of rotational symmetry and we have seen that there is no feed-back mechanism whereby the large field of the quadrupole mode could in turn serve as the primary for amplification of the dipole. In order to obtain feedback in this sense we

[7] H. Alfvén, M.N.R.A.S. **105**, 3, 383 (1945).
[8] L. Gurevich and A. Lebedinsky, J. Phys. U.S.S.R. **10**, 327, 425 (1946).

must make use of interactions without rotational symmetry.

1. There is a vast number of matrix elements involving tesseral harmonics, and we shall not try to undertake their classification and discussion. A simple relation between zonal and tesseral harmonics is obtained by a tilt of the rotational axis, according to (22). If in the amplification process (27) the fluid motion takes place about an axis inclined relative to that of the earth, as is apparently the case in the observed field, Eq. (27) will be transformed so that it involves elements of type

$$[\mathbf{T}_1^c \cdot \mathbf{S}_1 \times \mathbf{S}_2^c], \quad [\mathbf{T}_3^c \cdot \mathbf{S}_1 \times \mathbf{S}_2^c], \qquad (29)$$

and two others where the upper indices c are replaced by s.

The elements that provide feedback are of type (21). Limiting the permissible vectors to those indicated before, they are

$$[\mathbf{S}_2 \cdot \mathbf{T}_2^c \times \mathbf{T}_1^s], \quad [\mathbf{S}_2^c \cdot \mathbf{T}_2 \times \mathbf{T}_1^s], \qquad (30a, b)$$

$$[\mathbf{S}_2^c \cdot \mathbf{T}_2^s \times \mathbf{T}_1], \qquad (31)$$

and three others obtained by interchanging the upper indices. (Note the different selection rule for the upper indices as compared to (29).) Only the elements of type (31) give a component of the secondary magnetic dipole along the earth's axis. It follows that both the fluid motion $\mathbf{S}_2$ and the quadrupole field $\mathbf{T}_2$ must be inclined relative to the earth's axis. The same should then be the case for the velocity components, $\mathbf{T}_1$ and $\mathbf{T}_3$, producing the quadrupole field. But the two vectors in (31) have a shift in phase (i.e. in geographical longitude) of 90°; hence by (29) there should also be such a phase shift between the fluid components of the $\mathbf{T}$ and $\mathbf{S}$ types. In order to establish a concrete picture we may imagine that this whole complicated system precesses slowly about the earth's axis, as the inclined axes can hardly remain fixed. Now estimate the angle of inclination by means of the observed inclination of the dipole field, $\epsilon = 0.20$. Other things being equal, (31) is of the order ϵ^2 as compared to a rotationally symmetrical matrix element, hence quite small. An alternate feed-back mechanism derived from (30) is of the same order, and other feed-back mechanisms involving vectors that we have neglected here are of the order ϵ^3 or smaller. This type of feed-back mechanism is so complicated and artificial that it would hardly seem convincing.

2. *A more satisfactory feed-back mechanism* can be developed by taking into account the effect of turbulence upon the distribution of the magnetic field. It follows from general hydrodynamical principles that fluid motion in the core must be turbulent. From the viewpoint of the present theory the irregular features on any map of the magnetic field or its secular variation may be interpreted as the result of turbulent motion.

We shall first prove two useful theorems. Consider a surface of arbitrary shape, bounded by a contour C, located entirely inside the conducting fluid. We then get[9] from (9) and (10), on integration

$$(\partial/\partial t)\int \mathbf{B}_n dS = -\int \mathbf{E} \cdot d\mathbf{C}$$

$$= \int (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{C} - (\mu\sigma)^{-1} \int (\nabla \times \mathbf{B}) \cdot d\mathbf{C}.$$

Now, if the first integrand on the right-hand side is written $\mathbf{B} \cdot (d\mathbf{C} \times \mathbf{v})$, the integral can be given a simple geometrical meaning: It becomes $-\int \mathbf{B}_n dS$ where the integration extends over the strip that the contour C subtends in its motion during the time dt. Since $\int \mathbf{B}_n dS = 0$ for any closed surface, we find that

$$(d/dt)\int \mathbf{B}_n dS = (d/dt)\int \mathbf{A} \cdot d\mathbf{C} = O(\sigma^{-1}) \qquad (32)$$

where the term on the right-hand side tends to zero as the free decay effects are neglected. The operator (d/dt) has the usual significance, giving the rate of change when the surface and contour move with the fluid. By means of (32) we can understand why certain simple theories of inductive amplification do not succeed. A theory of such a type was proposed for the sun by Larmor.[10] Apply (32) to a circle $r = \text{const}$. Any rotationally symmetrical motion transforms such

[9] This theorem is due to T. G. Cowling who communicated it to us some months ago. The writer is particularly indebted for Dr. Cowling's generous permission to reproduce it here and to use the results in the context of this paper.

[10] J. Larmor, Brit. Assoc. Adv. Sci., Bournemouth Meeting, 1919, p. 159 (1920).

a circle into a family of similar circles, and eventually the circle will return arbitrarily close to its initial position. Now (32) asserts that, neglecting free decay, the flux threaded by the circle remains constant. This flux is produced by the modes of magnetic type, as may readily be verified. On admitting that free decay can, in the mean, only decrease the magnetic flux it follows that the magnetic modes cannot be amplified by a stationary flow of rotational symmetry. This is an alternate proof[9] of Cowling's earlier theorem.[11]

A second integral theorem is obtained by extending the integration over an arbitrary volume instead of over a surface. This gives, by (9) and (10)

$$(\partial/\partial t)\int \mathbf{B}dV = -\int \nabla \times \mathbf{E}dV = -\int \mathbf{n}\times\mathbf{E}dS$$

$$= \int \mathbf{n}\times(\nabla\times\mathbf{B})dS - (\mu\sigma)^{-1}\int \mathbf{n}\times(\nabla\times\mathbf{B})dS$$

$$= \int \mathbf{v}B_n dS - \int \mathbf{B}v_n dS - O(\sigma^{-1}), \qquad (33)$$

where dS is an element of the surface bounding the volume and $\mathbf{n}$ the normal to this surface, pointing outwards. Equation (33) can be rewritten as

$$(d/dt)\int \mathbf{B}dV = \int \mathbf{v}B_n dS - O(\sigma^{-1}), \qquad (34)$$

where we shall again neglect the second term on the right-hand side.

This relation may now be used to determine what happens to the field of a fluid particle during the motion of the latter. Consider, in turn, the three basic forms of displacement: translation, rotation, and deformation. For a pure translation, $\mathbf{v}$ can be taken out of the surface integral in (34) and the latter vanishes. Hence, if the fluid particle is displaced parallel to itself, it carries the field with it unchanged, both in magnitude and direction. Next, let the particle rotate about an arbitrary axis, $\boldsymbol{\omega}$ being the angular velocity and $\mathbf{r}$ the distance from a point on this axis. Then (34) gives

$$(d/dt)\int \mathbf{B}dV = \boldsymbol{\omega}\times\int \mathbf{r}B_n dS.$$

Now apply the vectorial identity

$$\int \mathbf{r}B_n dS = \int \mathbf{r}(\nabla\cdot\mathbf{B})dV + \int (\mathbf{B}\cdot\nabla)\mathbf{r}dV.$$

Using the particular properties of $\mathbf{r}$ and $\mathbf{B}$, we find that the first integrand vanishes and the second reduces to $\mathbf{B}$. Hence

$$(d/dt)\int \mathbf{B}dV = \boldsymbol{\omega}\times\int \mathbf{B}dV \qquad (35)$$

which shows that the magnetic field rotates together with the fluid particle.

In order to study deformation, let the volume be a cylinder with axis parallel to $\mathbf{B}$, bounded by two plane surface of area S perpendicular to the axis. Then, by (34)

$$(dt/dt)\int \mathbf{B}dV = (v_1+v_2)B\cdot S, \qquad (36)$$

where v_1 and v_2 are the outward velocities across the end faces of the cylinder. Relation (36) is the integral theorem corresponding to the differential equation (II, 6) and leads again to the results derived in Part II: Convergence of the motion in a plane perpendicular to the field produces amplification, divergence of the motion in this plane produces de-amplification.

It is a well-known fact that, with respect to a *scalar* property of the fluid, turbulence acts like a greatly enhanced diffusion. Now we know that, in the absence of motion, the field equation (II, 3) for $\mathbf{B}$ reduces to an equation of diffusion for the *vectorial* property $\mathbf{B}$, describing the phenomena of free decay, skin effect, etc., with a coefficient of diffusion equal to $(\mu\sigma)^{-1}$. From the preceding results we can infer that turbulent motion of the fluid acts so as to accelerate the free decay to a value that corresponds to the increased rate of diffusion measured by the coefficient of turbulent mixing. The formal analogy would seem to be as complete as in the scalar case. We shall not dwell here upon the mathematical aspects. The periods of free decay,

[11] T. G. Cowling, M.N.R.A.S. **94**, 39 (1934).

formerly found of the order of 20–50,000 years, may now be much smaller, of the order of the more rapid periods of the secular variation, say a thousand years or slightly less.

If the elements of turbulent flow rotate, the magnetic field vector rotates with them. For the electric quadrupole mode, **B** is perpendicular to the meridional plane. By rotating the fluid particles we get at once a component of **B** in this plane. In order that the *average* component in this plane may be finite, it is clearly necessary that the axes of the turbulent rotations be not entirely random in direction, but have some correlation either with the axis of the earth, or with the axis of the magnetic field. In order to visualize such a correlation in a very crude way, we may simply assume that a rotational component in the same sense as the earth's rotation is more probable than a rotational component in the opposite sense (or *vice versa*). If the magnetic field is small, the motion is controlled by the Coriolis forces, if it is large, the motion is controlled by the magneto-mechanical forces. It is therefore quite likely that the turbulent motion exhibits a correlation with the axis of the earth or with the axis of the electric quadrupole field, as the case may be. These qualitative considerations do not determine the sign of the correlation coefficient which must be assumed such that, in the mean, regenerative feedback results. We can now attribute the observed rapid change of the magnetic dipole terms, referred to above, to a turbulent feed-back mechanism of this kind rather than to a mean motion of the fluid. Since the observed dipole moment decreases, it appears that feedback is degenerative at the present time, at least in the surface layers. If the field maintains itself in the long run, the average feedback must, of course, be regenerative.

This concept of turbulent feedback can be expressed more precisely as follows. The rate of change of the dipole field is

$$dc_1/dt = [\mathbf{v} \times \mathbf{B} \cdot \mathbf{T}_1]. \qquad (37)$$

$\mathbf{T}_1$ has only a φ-component which may be written $a(r)\sin\vartheta$ where $a(r)$ is normalized. Hence

$$dc_1/dt = \int (\mathbf{v} \times \mathbf{B})_\varphi a(r) \sin\vartheta \, dV.$$

This form of the equations of motion brings out more clearly the significance of the matrix elements as coefficients of correlation for the turbulent motion and field. If **v** and **B** in (37) are developed in series of the orthogonal vectors, the right-hand side becomes a linear combination of matrix elements, as in (14), the first term being precisely the element (31). But there will now be an entire series of such elements involving higher harmonics. It should be pointed out that this feed-back effect, like any other, is a second-order effect in the sense of perturbation theory, but on purely physical grounds it would appear far more likely than the one sketched in the beginning of this section.

There is a lower limit to the size of the turbulent elements which can effectively contribute to changes of the field. By (I, 30) the lowest mode of free decay of a spherical particle is $\mu\sigma R^2/\pi^2$, or numerically, $0.13R^2$ when R is in meters. According to the evidence from the secular variation, significant changes of the field take place in periods of the order of a thousand years $= 3 \cdot 10^{10}$ sec. Equating to the decay period just mentioned, this gives $R = 500$ km. For turbulent elements of smaller size we can expect the effects of free decay to outweigh the turbulence so that the field is smoothed out before the correlation required for feedback becomes established. The large turbulent vortices, evidenced by the major traits of the secular variation, are several thousand km in diameter, and they will be quite effective in producing the phenomenon of feedback.

We are now in a position to estimate the magnitude of the electric quadrupole field. A statistical mechanism of feedback, as described, must involve a large amount of cancellation of random components and cannot be highly efficient. The magnitude of the dipole field in the upper strata of the core is 2–3 gauss; a value of 30 gauss for the quadrupole field would then keep the rate of feedback below 10 percent. The quadrupole field might of course be larger than this value. We shall, later on, derive an upper limit for the field.

Before leaving this topic, we might point to the unusual degree of statistical fluctuation that can be expected in a coupled magneto-mechanical system of this kind. Not all the fluctuations are

rooted in properties of the fluid motion. The theory of motional inductance might be said to be the theory of the differential Eq. (12). If one asks for the possible eigenvalues and eigenfunctions of this equation (for stationary $\mathbf{v}$) he will assume solutions of the form $A(r, \vartheta, \varphi) \exp(\gamma t)$, where γ is a complex constant. We know of one case where such a solution exists, and γ is purely imaginary. This occurs when the sphere rotates as if solid, and the field rotates with the sphere. But this is a trivial solution. The studies of the present writer have given little support to the hope that eigenvalues exist for more general types of stationary fluid motion. The operator appearing on the right-hand side of (12) is in general *skew*, being neither purely symmetrical nor purely antisymmetrical (see Part I). If a set of eigenvalues exist, they must therefore in general be complex numbers. The solution given by the amplifier mechanism (27) is not in the nature of an eigenfunction, since the primary field remains unchanged. If the presumption of the writer, namely, that the operator in (12) does not, in general, have eigenvalues, is correct, then one would expect fluctuations of the field of macroscopic magnitude, even if a system of feed-back amplification could be constructed on the basis of a stationary motion of the fluid. It seems questionable, however, whether the mechanism of feedback by turbulent correlation can be expressed in terms of strictly stationary motion; we have not investigated this point. We need hardly dwell on the familiar dynamical instability of fluid motion in large dimensions, giving rise to large-scale turbulence. One can therefore not be too surprised at the extraordinary amount of fluctuation, both in space and in time, which is exhibited by the observed field.

ENERGETICS

In the present theory the power required to maintain the field is much larger than one could have expected heretofore, from a consideration of the magnetic modes alone. An exhaustive, or even fairly complete analysis of the possible sources of power is beyond the scope of this paper. A satisfactory answer to the questions raised will no doubt be dependent upon a study of the features of the existing field and its variation. Inasmuch as the observed secular variation permits inferences about the character of the fluid motion, it might be possible to arrive at a decision as between the mechanisms outlined below.

If the mean field in the core is assumed to be 30 gauss, the energy density is 3.6 joule/m^3. This amount of energy must be supplied once during the time of free decay, say every 1000 years. Although we might expect large fluctuations in the magnitude of the field it is, in the absence of other information, perhaps best to assume that a field of this magnitude has in the mean existed throughout the geological history of the earth. If the matter in the core is taken to be mainly iron and its specific heat as $4R$ per mole (somewhat in excess of the Dulong-Petit value in order to allow for a contribution of the conductivity electrons at the high temperatures involved) the heat generated by the field is readily computed. Over a period of 5.10^8 years, comprising most of the known geological past, the rise in temperature of the core owing to this cause is 0.3 degrees.

We may now first assume that the power for the maintenance of the field is produced by thermal sources, for instance through radioactivity, inside the earth. It is possible to estimate the total heat required for this purpose. A hydrodynamical theorem states that hydrostatic equilibrium obtains when the surfaces of constant temperature coincide with the surfaces of constant mechanical (gravitational plus centrifugal) potential. Only the deviations, ΔT say, from the mean temperature of an equipotential surface can be utilized to generate motion. Since the fluid motion itself continually redistributes the heat throughout the core, ΔT can hardly amount to more than a few degrees, whatever the origin and shape of the variations. If an amount Q of heat is converted into energy of motion, the amount $QT/\Delta T$ must flow irreversibly from the regions of positive ΔT to the regions of negative ΔT, by the second law of thermodynamics. T is certainly of the order of several thousand degrees, so that $T/\Delta T$ should be at least 10^3; in practice the ratio of irreversible to reversible thermal effects is more likely to be about 10^4. Hence, the thermal sources must be of such magnitude as to be capable of heating the core by several hundred degrees in the course of $5 \cdot 10^8$ years.

Evidently this line of reasoning is not too attractive, even if it should prove possible to admit of radioactive sources of this magnitude. We shall not pursue the question further since a more plausible source of power for the fluid motions and the magnetic field can be found.

In this alternate view the power is a by-product of the change in the earth's speed of rotation caused by the lunar tide.[12] From astronomical observations

$$d\omega/dt = -2.5\cdot 10^{-22}\ \text{sec.}^{-2},$$

corresponding to a lengthening of the day by 1 second in 120,000 years. The accuracy of this figure may be estimated as about 10 percent. The angular momentum lost by the earth reappears in the orbital motion of the moon, as the latter gradually recedes from the earth. The power required for removal of the moon is small, and the existing theory[12] indicates that most of the kinetic energy of the earth's rotation is dissipated by tidal friction in the oceans. For the core alone, the average kinetic energy lost per unit volume is

$$[dE_{\text{kin}}/dt]_w = (\theta/V)\omega\cdot d\omega/dt, \qquad (38)$$

where V and θ are the volume and moment of inertia of the core. Numerically, $V = 1.75\cdot 10^{20}$ m^3 and $\theta = 8.4\cdot 10^{36}$ kg. m^2, obtained from the known density distribution.[13] Then (38) is equal to $2.7\cdot 10^{-2}$ joule/m^3 per year, an amount large enough to maintain an average field of 80 gauss against a mean decay period of 10^3 years. On the hypothesis that all the kinetic energy of the core instead of being dissipated by the tide, can be converted into magnetic energy, the figure may be interpreted as an upper limit for the magnitude of the internal field.

Since the retarding torque of tidal friction attacks the solid mangle of the earth, the central parts of the core will have a tendency to rotate faster than the mantle. In the stationary state there will be an increase of angular velocity with depth in the core, and the distribution of angular velocity will be such that angular momentum is carried from the inside out by frictional shear at a rate prescribed by the slowing down process. This condition of inhomogeneous rotation is a sufficient prerequisite for the functioning of the amplifier (27). Mathematically speaking, it is only one of two possible solutions. If the core is inhomogenous, as seems indicated by seismic observations,[14] the density increasing rather rapidly below a radius of $0.4R$, then the central part of the core can itself be subject to a rather large tidal deformation owing to the direct gravitational influence of the moon. If this tide slows down the central part of the core at a rate more rapid than that at which the solid mantle is slowed down by the oceanic tide, the velocity distribution would be the opposite of the one just described: the angular velocity would decrease with depth in the core and the transport of angular momentum by frictional shear would be towards the central part where the main tidal deceleration would take place. Without entering into details, it may be remarked that there are certain qualitative indications in the observed secular variation to the effect that the fluid core rotates somewhat slower than the solid body of the earth. Specifically, the local features of the magnetic field and the *foci* of the secular variation show a mean drift motion from east to west, well beyond the limits of observational errors.[15] Whether this effect can be interpreted along the lines sketched, could only be decided by an extensive study of the dynamics of these motions.

Next we can estimate the kinetic energy of the fluid motion. Let the velocity of a point in the core be written as $\boldsymbol{\omega}\times\mathbf{r}+\mathbf{v}$, where $\mathbf{v}=0$ if the core rotates synchronously with the solid mantle. If the fluid moves, the density of kinetic energy relative to the state of synchronous rotation is

$$\begin{aligned} E_{\text{kin}} &= (\rho/2)(\boldsymbol{\omega}\times\mathbf{r}+\mathbf{v})^2 - (\rho/2)(\boldsymbol{\omega}\times\mathbf{r})^2 \\ &= \rho(\mathbf{r}\times\mathbf{v})\cdot\boldsymbol{\omega} + (\rho/2)\mathbf{v}^2. \end{aligned} \qquad (39)$$

Assuming a typical value, $v = 0.01$ cm/sec. as inferred from the secular variation,[16] the second term on the right-hand side is about 10^{-4} joule/m^3, exceedingly small compared to the magnetic energy density. The first term, however, is of the order of 150 joule/m^3, and thus is ample. Now the first term is nothing but the angular momen-

[12] H. Jeffreys, *The Earth* (The Macmillan Company, New York, 1929), second edition, chapter 14.
[13] K. E. Bullen, Bull. Am. Seismol. Soc. **32**, 19 (1942).
[14] Reference 2, Appendix.
[15] See the maps by E. H. Vestine *et al.*, Carnegie Inst. of Washington Publ. 578 (1947).
[16] Reference 2, p. 209.

tum of the fluid particle about the earth's axis multiplied by ω; hence its integral over the sphere does not vanish when the motion is a zonal flow of type $\mathbf{T}_n$, as required by (27), and as provided by the mechanism of tidal deceleration just described.

This source of power for the maintenance of the magnetic field would appear to be characteristic of the earth's core; it cannot readily be generalized to apply to other celestial bodies. If it should be undertaken to explain the magnetic field of the sun by a mechanism of feed-back amplification, the inhomogeneous rotation must be produced by thermal effects and the energy required for amplification must ultimately be of thermal origin. Clearly, the thermal energy available in the sun is so vast that no direct comparison with the conditions in the earth's core is possible.

ELECTROMOTIVE FORCES

While the theory of feed-back amplification indicates how a magnetic field can be increased from small beginnings until the magneto-mechanical reaction prohibits further growth, it does not tell us how large the initial field can be. The theory would no doubt be more satisfactory if it could be demonstrated, on physical grounds, that such an initial field is likely to exist, and that it need not be too minute. We shall now proceed to show that electromotive forces acting in the core (for instance thermoelectric forces) can produce a small internal magnetic field of the electric quadrupole type.

According to the theory of the earth's figure the eccentricity of the core's boundary is nearly the same as that of the earth's surface.[13] It is therefore conceivable that a slight temperature difference exists between the polar and equatorial regions of this boundary, giving rise to a differential thermoelectric potential between its pole and its equator. As the electric conductivity of the solid mantle is likely to be much smaller than that of the core, this will not give rise to an appreciable current density, but merely to an electrostatic charge at the boundary. Now assume that a similar differential thermoelectric force exists at the internal transition layer or boundary located just below $r=0.4R$, which separates the iron from the heavier metals below.[14] For the sake of the present calculation, we shall idealize this by a well-defined boundary. The matter to both sides of the boundary will be assumed a good conductor; the differential e.m.f. then produces electric currents that flow from the pole to the equator in the upper stratum and from the equator to the pole in the lower stratum, or *vice versa*. Such a current system can be represented by a linear aggregate of vectors $\mathbf{S}_n$ (n even). We shall consider a single harmonic component of this field.

Let $r=r_1$ be the inner boundary, $r=R$ the boundary of the core, and let the three regions be distinguished by the indices (1), (2), and (3), counting from the inside out. The electric field is a vector of type $\mathbf{S}_n$, but in the static case it can as well be represented by

$$\mathbf{E}=\nabla\psi.$$

By means of (I, 10) we can then write for the magnetic field

$$\mathbf{B}=\mu\sigma(n+1)^{-1}(\mathbf{E}\times\mathbf{r}) \quad \text{for } \psi=r^n Y_n,(\vartheta,\varphi),$$

$$\mathbf{B}=\mu\sigma n^{-1}(\mathbf{E}\times\mathbf{r}) \quad \text{for } \psi=r^{-n-1}Y_n,(\vartheta,\varphi),$$

thereby fulfilling the field equation

$$\nabla\times\mathbf{B}=\mu\sigma\mathbf{E}.$$

In region (3) the magnetic field vanishes, since $\sigma=0$.

We now set in the three regions

$$\psi^{(1)}=ar^n Y_n,$$

$$\psi^{(2)}=(br^n+cr^{-n-1})Y_n,$$

$$\psi^{(3)}=dr^{-n-1}Y_n.$$

The boundary conditions at the inner surface require

$$\psi^{(2)}-\psi^{(1)}=\text{const.}\,Y_n=\phi_n,$$

where ϕ_n is the impressed electromotive potential. They require, moreover, continuity of $\mathbf{B}$, which also provides continuity of the radial component of the current. At the outer boundary $\mathbf{B}$ must vanish and the tangential components of $\mathbf{E}$ are continuous. We shall confine ourselves to writing down the solution for region (2). If σ_1 and σ_2 are the conductivities of regions (1)

and (2), we find

$$\psi^{(2)}=\frac{(n+1)(r/R)^n+(2n+1)(R/r)^{n+1}(\sigma_2/\sigma_1)}{(2n+1)(\sigma_2/\sigma_1)+(n+1)(1-\sigma_2/\sigma_1)(r_1/R)^{2n+1}}\left(\frac{r_1}{R}\right)^{n+1}\phi_n.$$

Now assume for simplicity $\sigma_1=\sigma_2$. If r_1 and r are not too close to R, we can neglect the first term in the numerator and the second term in the denominator, and get

$$\psi^{(2)}=(r_1/r)^{n+1}\phi_n$$

valid in the region just above the inner boundary. For the magnetic field of the quadrupole mode we find at $r=r_1$

$$\mathbf{B}=-(\mu\sigma/2)\sin\vartheta\cdot\phi.$$

If we put $\phi=1$ millivolt, $\sin\vartheta=\frac{1}{2}$, we get a field of about 3 gauss. Thermoelectric potentials are of the order of a few microvolts per degree temperature difference. Taking the acting e.m.f. to be 10 microvolts, we require an over-all amplification of about a thousand from the initial field to the final magnitude of the electric quadrupole mode.

APPENDIX

The following expressions for the matrix elements simplify and generalize those of Part I. Let $Z(r)$ designate a normalized radial eigenfunction. We define the following integrals where the prime denotes differentiation with respect to r.

$$F=R\int Z_\alpha Z_\beta Z_\gamma r dr,$$

$$G_\alpha=R^2\int (rZ_\alpha)'Z_\beta Z_\gamma dr,$$

$$H_\alpha=R^3\int Z_\alpha(rZ_\beta)'(rZ_\gamma)'r^{-1}dr.$$

We furthermore define integrals over the sphere

$$K=\int Y_\alpha Y_\beta Y_\gamma dS,$$

$$L_\alpha=\int\frac{Y_\alpha}{\sin\vartheta}\left[\frac{\partial Y_\beta}{\partial\vartheta}\frac{\partial Y_\gamma}{\partial\varphi}-\frac{\partial Y_\beta}{\partial\varphi}\frac{\partial Y_\gamma}{\partial\vartheta}\right]dS.$$

We may write this L, for short, since

$$L_\alpha=L_\beta=L_\gamma.$$

The last relation may be proved by means of integrations by parts. If now, as in Part I, the product of the normalization factors of the spherical harmonics is designated by N we obtain, by (I, 13) and (I, 14)

$$[\mathbf{S}(\alpha)\cdot\mathbf{T}(\beta)\times\mathbf{T}(\gamma)]=n_\alpha(n_\alpha+1)NFL,$$

$$[\mathbf{S}(\alpha)\cdot\mathbf{S}(\beta)\times\mathbf{T}(\gamma)]$$
$$=-\tfrac{1}{2}n_\alpha(n_\alpha+1)[n_\beta(n_\beta+1)+n_\gamma(n_\gamma+1)-n_\alpha(n_\alpha+1)]NG_\beta K$$
$$+\tfrac{1}{2}n_\beta(n_\beta+1)[n_\alpha(n_\alpha+1)+n_\gamma(n_\gamma+1)-n_\beta(n_\beta+1)]NG_\alpha K,$$

$$[\mathbf{S}(\alpha)\cdot\mathbf{S}(\beta)\times\mathbf{S}(\gamma)]$$
$$=[n_\alpha(n_\alpha+1)H_\alpha+n_\beta(n_\beta+1)H_\beta+n_\gamma(n_\gamma+1)H_\gamma]NL.$$

As a supplement to the tables in Part II, p. 207, we give below all the non-vanishing matrix elements containing dipole and quadrupole vectors of the type **S** only. To conform with the previous tables we use complex notation:

$$[\mathbf{S}_1\cdot\mathbf{S}_1{}^1\times\mathbf{S}_1{}^{-1}]=-4\pi i(2/3)N(H_\alpha+H_\beta+H_\gamma),$$
$$[\mathbf{S}_1\cdot\mathbf{S}_2{}^1\times\mathbf{S}_2{}^{-1}]=-4\pi i(2/5)N(H_\alpha+3H_\beta+3H_\gamma),$$
$$[\mathbf{S}_1\cdot\mathbf{S}_2{}^2\times\mathbf{S}_2{}^{-2}]=-4\pi i(4/5)N(H_\alpha+3H_\beta+3H_\gamma),$$
$$[\mathbf{S}_1{}^1\cdot\mathbf{S}_2\times\mathbf{S}_2{}^{-1}]=4\pi i(2/5)N(H_\alpha+3H_\beta+3H_\gamma),$$
$$[\mathbf{S}_1{}^1\cdot\mathbf{S}_2{}^1\times\mathbf{S}_2{}^{-2}]=4\pi i(2/5)N(H_\alpha+3H_\beta+3H_\gamma).$$

The Temperature Variation of the Thermal Diffusion Factors for Binary Mixtures of Hydrogen, Deuterium, and Helium

BYRON F. MURPHEY*
Department of Physics, University of Minnesota, Minneapolis, Minnesota
(Received August 1, 1947)

Experimental values of the thermal diffusion factor, α, have been measured by the twin bulb method for mixtures of H_2 and D_2, H_2 and He, and D_2 and He. Analyses were made by means of a mass spectrometer. The value of α for a mixture of H_2 and D_2 was found to increase slightly with temperature. For a mixture of H_2 and He the value of α was found to decrease slightly as the temperature is raised. The thermal diffusion factor was found to be measurable for a mixture of D_2 and He and also appeared to decrease with temperature.

I. INTRODUCTION

IN view of the large temperature variation of the thermal diffusion factor observed in the cases of isotopes of neon and argon by Stier,[1] it appeared worth while to make further observations in the cases of hydrogen and helium. The thermal diffusion factor, α, is related to the coefficient of diffusion, D, and the coefficient of thermal diffusion, D_t, by the equation

$$\frac{D}{D_t}=\alpha C_1C_2=k_t,$$

where C_1 and C_2 are the relative concentrations of the molecules and k_t is the thermal diffusion ratio. For a region in which there is a constant temperature gradient and in which the pressure is uniform, the following equation holds at equilibrium,

$$\text{grad}C_1=-k_t\cdot\frac{1}{T}\cdot\text{grad}T$$

where T is the absolute temperature. In the case of two bulbs together with a connecting tubing, where one bulb is at a temperature, T_1, and the other at a temperature T_2,

$$C_1{}^1-C_1{}^2=\alpha C_1C_2\ln\frac{T_2}{T_1}$$

where $C_1{}^1$ and $C_1{}^2$ are the observed concentrations of the heavier component, 1, in the bulbs at the temperatures T_1 and T_2, respectively.

Theoretical values of α have been calculated by Chapman[2] for cases in which the law of force between molecules is assumed to be a repulsive one of the type

$$F=\kappa r^{-\nu}.$$

Since this type of force law does not lead to any temperature dependence of α, evidently a more detailed account of the law of force is required, inasmuch as α is in many cases temperature dependent. One such more extensive calculation has been made by Jones[3] for the case

$$F=\kappa r^{-\nu}-\kappa' r^{-\nu'},$$

where $\nu=9$ and $\nu'=5$. This model does lead to a prediction of a strong temperature variation, but does not check in detail with experimental results. Such experimental evidence as obtained by Stier[1] and Grew[4] does indicate, however, that some such law of force is required in order to explain the observed temperature dependence of the thermal diffusion factor.

The thermal diffusion factors of mixtures of hydrogen and deuterium and of hydrogen and helium have been measured by Heath, Ibbs, and Wild,[5] by Grew,[6] and by Elliott and Masson.[7] In each case the method of analysis has been that of noting the variation in the thermal conductivity as the relative concentrations change. The experiments by Grew were designed to cover a wide range of temperatures, but in all cases the temperature of the conductivity cell was approximately room temperature. The conclu-

* National Institute of Health Junior Research Fellow.
[1] L. G. Stier, Phys. Rev. **62**, 548 (1942).
[2] S. Chapman, Proc. Roy. Soc. **A177**, 38 (1940).
[3] R. C. Jones, Phys. Rev. **59**, 1019 (1941).
[4] K. E. Grew, Proc. Roy. Soc. **A189**, 402 (1947).
[5] H. R. Heath, T. L. Ibbs, and N. E. Wild, Proc. Roy. Soc. **A178**, 380 (1941).
[6] K. E. Grew, Proc. Roy. Soc. **A178**, 390 (1941).
[7] G. A. Elliott and I. Masson, Proc. Roy. Soc. **A108**, 378 (1925).

sion reached by Grew was that the thermal diffusion factor for a hydrogen and deuterium mixture is not temperature dependent.

The purpose of the work to be described here was to repeat the measurements on hydrogen and deuterium, using a mass spectrometer for the purpose of analysis and to extend the measurements to include mixtures of hydrogen and helium and deuterium and helium.

II. APPARATUS AND PROCEDURE

The arrangement of the twin bulbs together with provisions for flowing gas from the bulbs into the mass spectrometer is illustrated in Fig. 1. Each bulb was maintained at the desired temperature either by means of a constant temperature bath, or by means of a cylindrical furnace. The gas leaks illustrated in Fig. 1 each consisted of a 20 in.-long piece of capillary copper tubing, 0.006 in. I.D., which was pinched near the point where the capillary tubing enters the valve. This type of gas leak has been described recently by Nier.[8] The valves permitted introduction into the mass spectrometer of gas from either of the two leaks. When the steel balls were so placed that the gas from one leak was fed into the mass spectrometer and the gas from the other leak was being pumped away, there was less than one percent contamination of the former by the latter.

Thus, merely by moving a pair of steel balls by means of magnets, it was possible to introduce gas from one or the other of the two bulbs into the mass spectrometer. Since the sampling of the gas in the bulbs took place continuously, it was possible to observe the approach to equilibrium. After equilibrium was attained, a set of eight analyses was made of the gas in each of the two bulbs. In each case, the experiment was repeated, interchanging the roles of the hot and cold bulbs.

The mass spectrometer which was used for analysis consisted of a specially constructed 60° apparatus with three collectors, so that ions of masses 2, 3, and 4 could be observed simultaneously. In addition, any pair of the outputs of the three feed-back amplifiers could be balanced against one another in order to provide direct reading of the ratio of the ion currents. A similar arrangement has been described recently by Nier, Ney, and Inghram.[9]

In the case of helium and deuterium the mass spectrometer described recently by Nier[8] was used, with the change that the collecting slit was made narrow enough to allow resolution of the helium and deuterium ion peaks, which differ in mass by one part in 160. In this case one peak was not balanced against the other, but rather a series of measurements of peak heights was taken on a self-balancing recording potentiometer.

The amount of separation obtained in the various experiments was large enough so that only a single run of the thermal diffusion process was necessary in order to obtain a measurable change in the relative concentration. Each of the values of α given in the tables is the result of four separate thermal diffusion experiments on the same gas sample. Since the values of α which are obtained are for various temperature

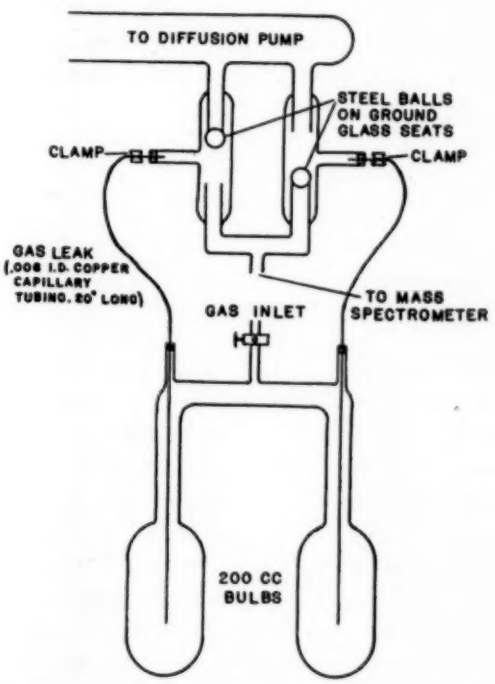


FIG. 1. Twin bulbs and gas leaks. The connecting tubing between the bulbs consisted of 25 cm of tubing of 1.13 cm² average cross section. The steel balls were $\frac{1}{32}$" in diameter. The pumping speed in the line marked "to diffusion pump" was about $\frac{1}{3}$ liter per second, whereas the pumping speed toward the mass spectrometer was about one percent of that value.

[8] A. O. Nier, Rev. Sci. Inst. **18**, 398 (1947).

[9] A. O. Nier, E. P. Ney, and M. G. Inghram, Rev. Sci. Inst. **18**, 294 (1947).

TABLE I. Experimental values of α for mixtures of hydrogen and deuterium.

Bulb temps.°K	T_r	% D_2	$C_1^1-C_1^2$	α	α (av.)	R_t (av.)
77–194	118	20.0	0.0194	0.1315		
77–194	118	19.3	0.0193	0.133	0.132	0.44
194–292	237	20.0	0.0095	0.146		
194–295	239	19.3	0.0091	0.140	0.143	0.48
194–295	239	9.1	0.0050	0.144		
273–359	312	20.0	0.0066	0.150		
273–361	314	19.3	0.0064	0.148	0.149	0.50

ranges, it is necessary to choose some average temperature as the one appropriate to the value of α which is measured. Brown[10] has shown that the proper temperature, T_r, for the temperature range T_1 and T_2 is given by the expression

$$T_r=\frac{T_1T_2}{T_2-T_1}\ln\frac{T_2}{T_1}.$$

The values listed under R_t are calculated by taking the ratio of the experimentally determined value of the thermal diffusion factor to the value that would be calculated on the basis that the molecules were hard spheres. The latter expression is

$$\alpha_{\text{hard spheres}}=\frac{105}{118}\frac{M_1-M_2}{M_1+M_2}.$$

The gases used in the experiments consisted of tank hydrogen (99.8 percent), tank helium (98.2 percent), spectroscopically pure helium, and a supply of 99 percent deuterium. In all cases the water vapor was frozen out with liquid air. Before introducing gases to the twin bulbs, the bulbs were torched while under a fore-pump vacuum. The gas mixtures were made by mechanical mixing with a Toeppler pump, so that diffusion alone would not have to be relied upon to bring about the same initial concentration in both bulbs. When checks were made to find the effect of small amounts of impurities, the helium was purified by passage through a liquid-air-cooled charcoal trap, and the hydrogen by passage through a palladium tube.

III. RESULTS

The experimental values of the thermal diffusion factor for mixtures of hydrogen and deuterium are presented in Table I. Extension of this series of measurements to higher temperatures was prevented due to the formation of HD, which may have been due to the catalytic action of the copper capillary tubing which was used for the leaks. It was possible to observe the growth of the HD^+ ion current at high temperatures, since the mass 3-ion current was at all times observable.

Some experiments were run on a sample of deuterium containing one percent of HD, over the temperature range 194° to 292°K. The average of α obtained was 0.061. Since in this case $\alpha_{hs}=0.127$, the value of R_t is 0.48. It is of interest to note that this value of R_t agrees closely with the corresponding value of R_t in Table I.

The results for hydrogen and helium mixtures are set down in Table II. The sample labeled 19.1 percent helium contained spectroscopically pure helium, whereas the other samples contained ordinary tank helium. In addition to the data presented in Table II, a separate series of measurements in which various bulb arrangements were tried yielded the value 0.120 for α. The temperature range for these runs was 194° to 287°K. One such arrangement consisted of one small bulb and one large bulb. In such a case nearly all the change in concentration appears in the small bulb. This was tried in order to be sure that the "dead" volumes were not affecting the measurements. Any such effect should have shown up in this case where one of the bulbs was much smaller than the other, and in fact smaller than the connecting tubing. In these test runs the method of sampling the gas was the same as that illustrated in Fig. 1.

The results for helium and deuterium are as follows: For a carefully purified sample in which

TABLE II. Experimental values of α for mixture of hydrogen and helium.

Bulb temps.°K	T_r	% He	$C_1^1-C_1^2$	α	α (av.)	R_t (av.)
77–194	118	19.1	0.0191	0.133		
77–194	118	19.2	0.0189	0.131	0.132	0.44
194–287	235	19.1	0.0073	0.121		
194–287	235	19.2	0.0073	0.121	0.121	0.41
273–360	314	19.1	0.0053	0.123		
273–362	315	19.2	0.0053	0.122	0.122	0.41
273–476	357	17.9	0.00965	0.117		
273–479	358	20.3	0.0104	0.114	0.116	0.39

[10] H. Brown, Phys. Rev. **58**, 661 (1940).

the deuterium was passed through a palladium leak and in which the helium was purified by passage through a liquid-air-cooled charcoal trap, the value of α was found to be 0.027 ± 0.003. The change in concentration was 0.00174, and the initial concentration of the deuterium was 0.213. The temperature range was 194° to 286°K. Previous to this analysis, a brief series of runs on a sample which contained about two percent of impurities yielded the values 0.033, 0.024, and 0.020 for the temperature ranges, 77° to 194°, 194° to 291°, and 273° to 441°K. These measurements are considered to be accurate to within 15 percent, whereas the values of the thermal diffusion factors given in Tables I and II are consistent within two percent.

IV. DISCUSSION

In the case of a mixture of hydrogen and deuterium there is little doubt but that the measurements indicate a real, though small, variation of α with temperature. The variation of α shown for a mixture of hydrogen and helium also appears to be larger than the experimental error involved. It is interesting to note that over the range of temperatures under investigation the value of α appears to decrease as the temperature is raised.

The case of deuterium and helium mixtures is of particular interest because the thermal diffusion factor which has been observed here cannot arise from the difference in mass, but must rather be due to the difference in diameters. That is to say, α depends principally upon the law of force. For this reason one would expect the temperature dependence to show up to a greater degree. Such appears to be the case. Since the value of the thermal diffusion factor decreases with increasing temperature, one may think of the deuterium as increasing in hardness at a more rapid rate than does the helium. In terms of diameters of the molecules, this would mean a greater difference in diameters at low temperatures and hence a greater value of α at low temperatures.

The only comparison with theoretical values of α which can be made at present is for the simple repulsive inverse power model for which Chapman[2] has given a rather complete discussion. Following the method of calculation outlined in his article, values of the thermal diffusion factor have been calculated for the two cases in which the inverse power is eleven and thirteen. For purposes of comparison, the values of the thermal diffusion factors obtained in the present and earlier work and those obtained by calculation from the inverse power model are given in Table III. The values of α given from earlier work have been calculated from the published values of k_t and the concentrations, using the relationship

$$k_t = \alpha C_1 C_2.$$

TABLE III. Experimental and theoretical values of α.

Mixture	α Present work	α Previous work	Calculated (1st approx.) $\nu=11$	$\nu=13$
H_2-D_2	0.149	0.170[6] (av.)	0.148	0.164
H_2-He	0.122	0.178[5] (80% H_2)	0.122	0.135
		0.147[5] (90% H_2)		
		0.146[7] (60% H_2)		
$He-D_2$	0.024	None	0.024	0.027

No explanation has been found for the discrepancies in the experimental results. The significant features of the comparison are that the α for a mixture of hydrogen and deuterium is somewhat larger than that for hydrogen and helium mixtures both in theory and as a result of experiment, and that the observed value for a mixture of helium and deuterium is about what one would expect from calculations. It should be noted that α in the latter case is positive, that is, the heavier and larger molecule (deuterium) is concentrated in the cold bulb. No stress should be placed on the over-all agreement of the experimental values with the theoretical values for $\nu=11$, since the very existence of the temperature effect points to the inadequacy of the simple inverse power model.

The author takes pleasure in acknowledging the many helpful discussions with Professor Alfred O. Nier. Thanks are also due Mr. R. Thorness for his part in the construction of the mass spectrometer. The research was supported by a grant from the Research Corporation.

PHYSICAL REVIEW VOLUME 72, NUMBER 9 NOVEMBER 1, 1947

The Theory of Liquid Helium*

LASZLO TISZA
Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received May 17, 1947)

A quasi-thermodynamic theory of liquid helium is developed based on very general assumptions regarding the energy spectrum of the liquid. Below the λ-point helium can be considered as a mixture of two fluids capable of two velocity fields. The macroscopic hydrodynamic equations of this system are derived. These are in the same relation to ordinary hydrodynamics as a two-body is to a one-body problem in point mechanics. The flow connected with the center of mass coordinate is associated with the transport of mass at constant entropy, the one connected with the relative coordinate (internal convection) is associated with an entropy transport at constant density. The results are in excellent agreement with experiment, in particular with the measurements of the temperature waves (second sound), the existence of which was predicted by the theory. The Bose-Einstein liquid introduced earlier is a possible molecular model consistent with the macroscopic results. Landau's theory is discussed.

I. INTRODUCTION

SEVERAL years ago F. London[1] brought forward evidence to support the idea that the peculiar phase transition of liquid helium at 2.19°K (λ-point) might be regarded as caused by the condensation mechanism characteristic of the ideal Bose-Einstein gas distorted by the presence of molecular forces which of course cannot be ignored in the case of a liquid. He showed that the various earlier attempts to explain the λ-transition by the familiar mechanisms of phase transitions of the second kind (order-disorder transition in ordinary space) are incompatible with the actual van der Waals forces of helium and with the requirements of quantum mechanics. In this situation the example of an order in momentum space, as presented by the Bose-Einstein condensation, seemed to offer a very welcome and suggestive new possibility. Moreover, the Bose-Einstein condensation demonstrated a very peculiar feature. It led to a kind of phase equilibrium of two phases inter-penetrating in ordinary space but separated in momentum space. There were good reasons to suspect that the "condensed" fraction of the atoms was the seat of a superfluid mass transfer.

Employing these ideas the present author was able to show[2] that the kinetic effects to be expected in an ideal Bose-Einstein gas revealed a close similarity to the well-known effects in helium II. Nevertheless, the theory did not seem very convincing since it was difficult to understand how the properties of a liquid could be interpreted, even qualitatively, in terms of a gas theory. The rigorous treatment of this problem would require the solution of the quantum-mechanical many-body problem. The mathematical difficulties involved have proved so far to be so tremendous that as yet it has been impossible to substantiate these ideas by the development of a rigorous molecular theory.

In view of this situation, the author has tried to avoid these difficulties by developing a somewhat more phenomenological theory of the various thermo-hydrodynamic effects in helium II.[3,4] (See also H. London.[5] For an account of the experiments, we refer to Keesom's monograph on helium.[6])

There are good reasons to expect that a quasi-thermodynamic method should be adequate to handle the problem of helium II. The peculiar properties of helium II are usually described by

* The research reported in this paper was made possible in part through support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army Signal Corps, the Navy Department (Office of Naval Research), and the Army Air Forces (Air Material Command), under the Signal Corps Contract No. W-36-039 sc-32037.

[1] F. London, Nature **141**, 643 (1938); Phys. Rev. **54**, 947 (1938).

[2] L. Tisza, Nature **141**, 913 (1938).

[3] L. Tisza, C. R. Paris **207**, 1035 and 1186 (1938).

[4] L. Tisza, J. de phys. et rad. **1**, 165 and 350 (1940).

[5] H. London, Proc. Roy. Soc. **A171**, 484 (1939).

[6] W. H. Keesom, *Helium* (Elsevier Publishing Company, Inc., Amsterdam, 1942).

the statement that this substance is superfluid and super-heat-conducting. Actually it is important to realize that the behavior of helium II cannot be characterized by assuming extremely small or large values of *kinetic coefficients* such as viscosity and heat conductivity. It is rather that the usual *differential equations* governing capillary flow and the transfer of heat have to be replaced by equations adapted to the particular mechanisms effective in helium II. This change in the differential equations rather than the assumption of an extreme value of the conductivity is characteristic also of superconductivity, as is apparent from the well-known theory of F. and H. London. While kinetic coefficients depend very sensitively on the nature of the molecular forces, this is usually not the case for the differential equations themselves.

The main point of the theory was the definition of the "Bose-Einstein liquid," uniting some properties of a liquid and some of a Bose-Einstein gas in a self-consistent scheme. Whereas the discussion of such a "hybrid" system could not be carried out with quantum-mechanical methods, it proved to be easily manageable from the much more schematic quasi-thermodynamic point of view. In particular, the notion of "condensation in momentum space" was extended to the Bose-Einstein liquid. Hence, below the condensation temperature, this system was supposed to have a heterogeneous character implying the existence of *two* velocity fields. The existence of two fields rather than one resulted in an "internal convection" carrying energy and entropy, but not associated with any net transfer of matter. Hence the macroscopic hydrodynamics of helium II proved to be of greater complexity than that valid for other substances and allowed the correlation of phenomena which appeared paradoxical from the point of view of ordinary hydrodynamics. The theory led also to various predictions which have been subsequently verified by experiment. The most important of these was the conclusion that inhomogeneities of the temperature would propagate according to a *wave equation* rather than the usual *parabolic equation* of heat conduction. The velocity of propagation of these "*temperature waves*" was computed as a function of the temperature. It was also indicated that an experimental verification could be obtained by exciting the temperature waves through periodic heating of helium II. A short report of these results appeared in 1938 in the form of preliminary notes,[3] but the detailed account[4] was, because of wartime conditions, not generally available until recently.

Landau, apparently unaware of these results, criticized the Bose-Einstein theory on the basis of the earlier note[2] and advanced instead a theory in which he maintained the assumption of the heterogeneous nature of helium II, but tried to justify it from a different molecular point of view based on quantum hydrodynamics.[7] As soon as the possibility of the two velocity fields was granted, the development of the macroscopic hydrodynamic equations had to proceed on the same lines as those based on the Bose-Einstein theory. In particular, Landau concluded that there should be two different modes of sound propagation in helium II with different velocities for the "first" and "second" sound. Landau did not discuss the physical meaning of this second sound wave. According to the recent account of Peshkov,[8] it was only after an attempt to excite the second sound by mechanical means had failed that the question was reconsidered by Lifshitz,[9] who concluded that the excitation of second sound could best be obtained by periodic heating of the liquid. Thus the second sound proved to be identical to the temperature waves predicted earlier on the basis of the Bose-Einstein theory. The existence of these temperature waves has been experimentally demonstrated by Peshkov[8] who measured also their velocity of propagation as a function of temperature. These results have been recently confirmed by an interesting indirect method suggested by Onsager.[10] The measurements of the temperature dependence of the velocity of propagation of second sound confirm the present author's prediction rather than that of Landau.[7,9] It may be noted that the theoretical curves do not contain any adjustable constants but are

[7] L. Landau, J. Phys. U.S.S.R. **5**, 71 (1941).
[8] V. Peshkov, J. Phys. U.S.S.R. **10**, 389 (1946); *ibid.* **8**, 381 (1944).
[9] E. Lifshitz, J. Phys. U.S.S.R. **8**, 110 (1944).
[10] C. T. Lane, H. Fairbank, H. Schultz, and W. Fairbank, Phys. Rev. **70**, 431 (1946); *ibid.* **71**, 600 (1947). My thanks are due Professor Lane and collaborators for communicating their results before publication.

evaluated with the empirical values of the entropy measured previously by Kapitza.[11]

The aim of the present paper is to clarify the relation between the theories and experiment, whereby it seems advantageous to proceed as far along phenomenological lines as possible. Such a procedure was adopted previously in reference 4, but the present paper is a further step in this direction. Thus it will be possible to avoid the pitfall of justifying any special molecular model by experimental results, which in reality can be also understood under broader assumptions.

Our procedure will then be to advance certain general assumptions regarding the energy spectrum of liquid helium, expressing the fact that this system unites in a peculiar way some characteristics of a liquid and of a gas (Section II). The lowest state of the system has liquid characteristics. The excited states can be classified into *compressional modes* which are also liquid-like (Debye phonons) and *shear modes* corresponding to the rearrangements of the molecules at constant density which are supposed to have a gas-like character. The technique of drawing conclusions from these assumptions will be essentially thermodynamic (Sections III and IV). Such an extension of classical thermodynamics seems to be an adequate method of dealing with second-order transitions. A discussion from the molecular point of view is, of course, of the greatest interest and will be outlined in Section V so far as it seems possible at present. Landau's theory will be discussed in Section VI.

II. THE FUNDAMENTAL ASSUMPTIONS

Let us consider a macroscopic system of helium atoms enclosed in a box of volume V. The characteristic values of the Schroedinger equation of the system, in particular the lowest energy E_0, are determined by the properties of the helium atoms and in addition depend on the volume V as a parameter. Two assumptions will now be made regarding the properties of this lowest state of the system. After a short discussion of the meaning of these assumptions, two others will follow regarding the excited states. Postulate (*a*). *The function $E_0(V)$ has a minimum for some value $V=V_0$ of the volume:*

$$(dE_0/dV)^{V=V_0}=0, \quad (d^2E_0/dV^2)^{V=V_0}>0. \quad (1)$$

[11] P. Kapitza, J. Phys. U.S.S.R. **5**, 59 (1941).

Postulate (*b*). *The state $E_0(V)$ is characterized by a vanishing microscopic rigidity.*

Assumption (*a*) expresses the fact that at the temperature of absolute zero and with a vanishing external pressure $P=-dE_0/dV=0$, the system is in a condensed, and not in a gaseous, state. Assumption (*b*) specifies further that this state is liquid rather than solid. The postulate that a system should have a vanishing coefficient of rigidity at *absolute zero* is very far reaching. Indeed, according to classical statistics, one would expect every system which is in thermodynamic equilibrium at absolute zero to be in a crystalline state. Hence it is reasonable to define a system satisfying postulates (*a*) and (*b*) as a *quantum liquid*. Significant conclusions regarding the properties of such systems can be obtained from an analysis of the notion of "microscopic rigidity." This terminology was introduced[12,13] in order to account for the fact that ordinary (classical) liquids are in some way more reminiscent of solids than of dense gases. Even in liquids there are potential barriers opposing the rearrangements of the molecules at constant density. (Such rearrangements will be of importance for this theory; we shall refer to them briefly as *isopycnic rearrangements*.) The potential barriers constantly collapse under the impact of thermal agitation, hence no macroscopic rigidity results. There is, however, a so-called microscopic rigidity which manifests itself in various visco-elastic effects.

If a shearing stress is applied at a high frequency, the system will react as a solid and propagate transverse elastic waves. For liquids of low viscosity the frequency required for such a behavior is in the thermal range (hypersonic in contrast to ultrasonic waves). The specific heats of monatomic liquids near the freezing point show good evidence for their presence.[14]

At low frequencies, the microscopic rigidity gives rise to viscosity. This *liquid-type* viscosity should be distinguished from the *gas type* or *transport* viscosity well-known from the kinetic theory of gasses; we will call them *dynamic* and

[12] L. Brillouin, *Les tenseurs en mécanique et en élasticité* (Masson et Cie, Paris, 1938).

[13] J. Frenkel, *Kinetic Theory of Liquids* (Clarendon Press, Oxford, 1946).

[14] E. Bauer, M. Magat, and M. Surdin, J. de phys. et rad. **7**, 441 (1937). Also p. 349 of reference 11.

kinetic viscosity, respectively. The two viscosities give rise to the same type of macroscopic effects, but they can be clearly identified from their temperature dependences.

For the kinetic viscosity one has

$$\eta_{\text{kin}} \sim \rho \bar{c} \bar{l}, \tag{2}$$

where ρ is the density, $\bar{c}$ the mean thermal velocity, and $\bar{l}$ the mean free path. η_{kin} increases slowly with the temperature.

For the dynamic viscosity one has a strongly negative temperature coefficient, since

$$\eta_{\text{dyn}} \sim e^{A/kT} \tag{3}$$

where A is an activation energy clearly indicating the presence of potential barriers. The situation is illustrated in Fig. 1 where the dynamic viscosity of liquid hydrogen is compared with the kinetic viscosity of hydrogen gas. The case of hydrogen is typical of classical liquids.

The most spectacular effect of the *microscopic* rigidity is *freezing*. As the thermal agitation decreases with temperature, the potential barriers opposing molecular rearrangements organize into a crystalline pattern and give rise to the *macroscopic* rigidity of the solid.

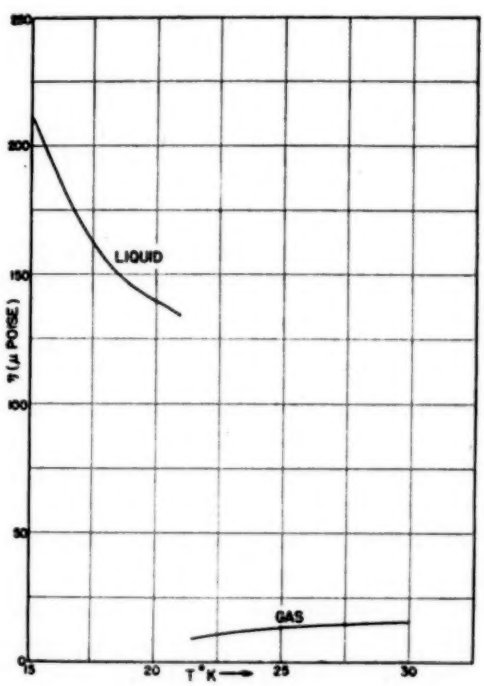


FIG. 1. Viscosity of liquid and gaseous hydrogen, from W. H. Keesom and G. E. MacWood, Physica **5**, 745 (1938); A. Van Itterbeck and A. Claes, Physica **5**, 938 (1938).

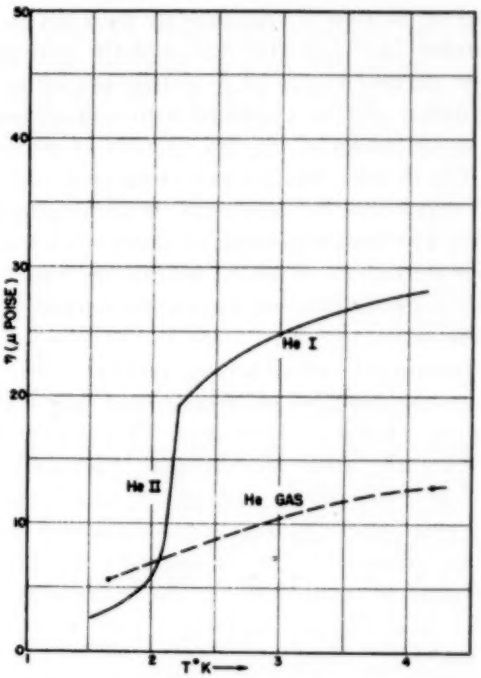


FIG. 2. Viscosity of liquid and gaseous helium from pp. 106 and 267–8 of reference 5.

The important point for the present purpose is that the propagation of transverse hypersonic waves, dynamic viscosity, and freezing are all due to the same factor, the microscopic rigidity. As helium does not solidify at low pressures, one concludes that the microscopic rigidity vanishes, as was postulated in (*b*). According to the above discussion this implies also the vanishing of the dynamic viscosity and the absence of transverse hypersonic waves. The first conclusion finds a striking verification in the measurement of viscosity. Figure 2, which shows the viscosity of helium, is to be compared with Fig. 1. Attention is drawn to the viscosity of helium gas and liquid helium I, since in the case of helium II additional complications arise which will be discussed below. It should be pointed out that in our terminology it is helium, not just helium II, that is a quantum liquid. This terminology is at variance with the rather generally adopted view that helium I is "classical" or a "normal" liquid.

A qualitative discussion of postulate (*b*) from the point of view of quantum mechanics will be

found in Section V. At present we turn to the discussion of the excited states of the system.

The excited states of a continuum or quasi-continuum can be classified into compressional and shear modes of motion. (P. 277 of reference 12.) The density fluctuations connected with the first type can be analyzed into longitudinal waves. The elastic quanta, or phonons, obtained by quantization of these waves are the "elementary excitations" of the compressional modes of motion.

Regarding the shear modes, postulate (*b*) leads only to the negative statement that they cannot have the character of waves. This gap will be filled by the next assumption: Postulate (*c*). *The shear modes of motion of liquid helium (isopycnic rearrangements of the molecules) have a gaseous character. The "elementary excitations" correspond to translations of atoms (or groups of atoms) with definite values of mass momentum and energy.*

In accordance with the program formulated in the introduction, postulate (*c*) is kept in very general terms, since the macroscopic equations to be derived in the next two sections do not depend on the molecular interpretation of the "gas" introduced in (*c*). A molecular interpretation is, of course, important for the intuitive understanding of the theory and will be discussed in Sections V and VI. At present it will suffice to mention that a molecular model compatible with postulate (*c*) is the Bose-Einstein liquid. In this case the "elementary excitations" correspond to helium atoms in translational Bloch-type states.[1,4] In reference 4 the term "translational quanta" was used to emphasize that only the translational energy in excess over the zero-point energy is considered to be the energy of the "molecule."

Also a somewhat different molecular model could possibly be considered. In this case the "elementary excitations" of postulate (*c*) are groups of atoms possessing, besides their linear momentum, also an internal angular momentum. Thus, the monatomic "molecules" of the Bose-Einstein liquid are replaced by polyatomic "molecules." This difference is irrelevant for most of the thermo- and hydrodynamic phenomena considered in this paper.

The most important property of the Bose-Einstein liquid is its "condensation" in momentum space.[1,4] The essential features of this condensation phenomenon will now be postulated for our more general case.

While the elementary excitations introduced in (*c*) have the main attributes of molecules in the kinetic theory of gases (constant mass, momentum, and energy) this is true only if a definite quantum state is considered. At absolute zero no "molecules" are present, but they can be created by means of thermal excitation. In every quantum state one can distinguish the total mass of the corresponding "molecules" from the total mass of the system. On averaging over a canonical ensemble, the same kind of subdivision is obtained as a function of temperature. Hence the density of the system will be subdivided into two parts:

$$\rho=\rho_n+\rho_s \tag{4}$$

where ρ_n is the density connected with the "molecules" of the gas and ρ_s refers to the "background" in which the molecules are moving. The subscripts refer to "normal" and "superfluid," a terminology which will be explained below. The ratio ρ_n/ρ is a very important property of the system. The last postulate of the theory will be concerned with it: Postulate (*d*). *There exists a temperature T_0 such that*

$$\rho_n/\rho\begin{cases}=0 & \text{if } T=0 & \text{He II}\\ <1 & 0<T<T_0 & \text{He II}\\ =1 & T\geqq T_0 & \text{He I}\end{cases}$$

with ρ_n/ρ monotonically increasing between $T=0$ and T_0. T_0 is to be identified with the lambda-point of helium.[15]

Let us consider now a volume element in helium II which is small compared to macroscopic dimensions but big enough to contain not only a large number of helium atoms but also a large number of "molecules." Averaging over the translational motion of the "molecules," one obtains a drift velocity $\mathbf{v}_n$ which will not in general coincide with the velocity of the "background" $\mathbf{v}_s$. In contrast with ordinary liquids and also with helium I, one has two densities and correspondingly two velocity fields. The

[15] Actually ρ_n/ρ will also be a function of the pressure P and this would have to be taken into account in a refinement of the present theory.

total mass current density is

$$\mathbf{j}=\rho_n\mathbf{v}_n+\rho_s\mathbf{v}_s. \tag{5}$$

According to (*c*) every "molecule" has a momentum. Averaging over this momentum current one gets, through the procedure well-known from the kinetic theory of gases, a stress tensor. The diagonal elements will be equal to a pressure P_n (*which is to be distinguished from the liquid pressure* P) while the non-diagonal elements represent a viscous stress. This is, of course, of the kinetic type and is uniquely associated with the flow $\mathbf{v}_n$. No kinetic viscosity is associated with the flow of the background $\mathbf{v}_s$. Since the dynamic viscosity was found to be totally absent in helium, none of the mechanisms of viscosity is effective for the flow $\mathbf{v}_s$ which may hence be called *superfluid*. In particular at absolute zero at which $\rho_n=0$, $\rho_s=\rho$, helium II as a whole is superfluid.

The character of the two currents will be essentially different; this is reflected most clearly in the boundary conditions for the tangential and perpendicular components of the velocities at a solid wall at rest:

$$v_{n\parallel}=0, \quad v_{n\perp}=0 \quad \text{whereas only} \quad v_{s\perp}=0. \tag{6}$$

The kinetic viscosity of the gaseous component entails the "normal" boundary condition for the velocity, while no condition is obtained for the tangential component of the superfluid flow. The terminology "normal" and "superfluid" is thus explained.

The superfluid component can "slip" along a solid surface which is its most essential characteristic, more so than the absence of dissipation. Actually, according to experimental evidence, the superfluid flow may be associated with a dissipation of the turbulent type as soon as a critical velocity is exceeded.

The slip of the superfluid component of helium II is most effectively demonstrated in the well-known creeping of the Rollin film.[6] This is, however, a complex phenomenon for quantitative discussion, corresponding to the fact that the general hydrodynamic equations of helium II are most unwieldy. It is satisfactory that the boundary conditions (6) lead to some simple conclusions which can be quantitatively checked by experiment. This will be seen in the next section.

Finally, we have to consider the specific entropy s of the system. According to our classification of the excited states, s consists of two terms

$$s=s_n+s_{\text{phon}} \tag{7}$$

corresponding to the shear and compressional modes of motion. As indicated by the subscript, s_n is associated with the *normal component* ρ_n of the liquid while the contribution s_{phon} of the Debye phonons is associated with the *liquid as a whole*.

III. HELIUM II AS A MIXTURE

The proper understanding of the heterogeneous nature of helium II is made somewhat difficult by the fact that it does not conform to the terminology of the classical phase rule with its clean cut distinction between "phase" and "independent component." In one connection it is best to consider helium II as a two-phase system and in another rather as a two-component system. Both analogies have to be handled with certain qualifications.

The two-phase point of view has been emphasized by London[1] in connection with the ideal Bose-Einstein gas. Below the condensation temperature this system can be considered as a gas in equilibrium with a condensed phase, the phase separation taking place in momentum space rather than in ordinary space. The saturation gas pressure $P_n=P_n(T)$ is a function of the temperature alone. In the liquid the situation is complicated by the fact that in addition to the gas pressure P_n one has liquid pressure P. The relation between P and P_n can be better understood in terms of the second analogy: the pressure P_n appears then as an "osmotic pressure" of the normal component dissolved in the superfluid. Because of the existence of semi-permeable membranes, P_n is subject to direct observation and this analogy will serve us as a useful guide throughout this section. Two qualifying remarks have to be observed, however. In the theory of solutions the total amount of each component is constant. In helium II this is true only if adiabatic conditions are maintained. In case of heat transfer there is a transition of one component into the other which has no analogy in the case of ordinary solutions. This requirement of maintenance of adiabatic conditions is of im-

portance also from the experimental point of view.

There is a second point in which P_n differs from the conventional osmotic pressure. A gradient of the osmotic pressure in a salt solution, for instance, gives rise to diffusion (second-order effect). In virtue of postulate (c), however, a gradient in P_n gives rise to a current (first-order effect) i.e., in this connection P_n has the character of a gas pressure. This idea will be developed in the next section.

The simplest and most direct experimental method for the verification of the heterogeneous nature of helium II and of the boundary conditions (6) consists in the measurement of the moment of inertia of a rotating vessel containing helium II. As the superfluid component will not take part in the rotation, the apparent moment of inertia of the liquid will be

$$I = I_0 \rho_n / \rho,$$

where I_0 is the moment of inertia above the λ-point. This effect was first predicted by Landau,[7] and the decrease of the moment of inertia was experimentally demonstrated by Andronikashvilli.[16] This experiment constitutes the most direct, although not the most accurate determination of the important quantity ρ_n/ρ. The experimental results can be approximately represented in the form

$$\rho_n/\rho = (T/T_0)^r, \tag{8}$$

where r is a constant of value about 6. More precise methods discussed below give $r = 5.5$. It should be emphasized that *this is a purely empirical relation which cannot be derived within the framework of the quasi-thermodynamic theory.*

A related, though somewhat more involved, case is the measurement of the viscosity by means of the oscillating disk method. Although the discussion goes somewhat beyond the scope of the quasi-thermodynamic method, it should be given here because of the special interest of the problem.

As in the previous case, the presence of the superfluid component can be ignored. The results can be understood in terms of the elementary formula (2), valid for the kinetic viscosity provided the density ρ is replaced by ρ_n. In order to obtain the temperature dependence of η, one has to realize that $\rho_n l = \text{const}$. This is equivalent to the statement that the "molecules" of the normal liquid collide only with each other. The justification and meaning of this assumption can be discussed only from the molecular point of view. (Compare Section V.) It should be mentioned also that in reference 4, the author has used at this point the ad hoc assumption $\rho l = \text{const.}$, which is in conflict with the fundamental assumptions of the theory. This was pointed out to the author by Professor Onsager to whom he is indebted also for the correct interpretation of the experiments. The earlier discussion of the viscosity measurements has also been rightly criticized by Keesom.[6] The present assumption leads to the temperature dependence $\eta \sim \tilde{c} \sim T^{\frac{1}{2}}$ for both helium I and II. This result seems to conflict with the measurements of Keesom and MacWood[17] reproduced in Fig. 2, according to which the temperature dependence of the viscosity of helium II is essentially given by $\eta \sim \rho_n/\rho = (T/T_0)^r$. An analysis of the experiment reveals, however (compare formula (1) in reference 17), that the quantity actually measured is the *viscosity multiplied by the density* which is $\eta\rho_n$ in the case of helium II. Keesom and MacWood have evaluated their measurement under the assumption of having obtained $\eta\rho$. Hence the curve in Fig. 2 has to be corrected by multiplication with the factor ρ/ρ_n, which is in agreement with the above result of the theory. Landau's interpretation of the viscosity measurement[18] is in line with the present discussion.

The next special case to be considered is that of a very thin capillary in which the flow velocity of the normal component is negligibly small. Such a capillary acts as a semi-permeable membrane, being permeable to the superfluid component alone; it can be used for the measurement of P_n in much the same way as the osmotic pressure of a salt solution is measured. The analogy with the osmotic cell leads to a qualitative understanding of the well-known fountain phenomenon in helium II.[6]

[16] E. Andronikashvilli, J. Phys. U.S.S.R. **10**, 201 (1946).

[17] W. H. Keesom and G. E. MacWood, Physica **5**, 737 (1938); The theory of the measurement is given in G. E. MacWood, Physica **5**, 374 (1938).

[18] L. Landau, J. Phys. U.S.S.R. **8**, 1 (1944).

We proceed now to investigate the "osmotic cell" with quantitative thermodynamic methods. In order to avoid complications connected with the liquid-vapor interface (evaporation and consequently non-adiabatic transition of the normal component to the superfluid state) an osmotic cell will be constructed as follows: A small part of a long cylinder filled with helium II is compressed adiabatically by means of a semi-permeable piston. (This can actually be realized using a porous material.) During the compression the superfluid component flows across the piston, thus keeping the total density ρ constant while the values of ρ_n, P_n, and consequently also of the temperature T are increased over their values in the helium bath.[19] *Since ρ is constant such compressions will be called isopycnic.*

The basis for the application of thermodynamics to the isopycnic compression is the assumption that the flow across the piston *carries along no entropy and is reversible.* This assumption was first advanced by the author[2] who pointed out also its approximate nature.[4] In fact, as indicated above, the phonon contribution to the entropy s_{phon} is associated with the whole liquid and is carried along in particular also with the superfluid component. Hence the above assumption means that

$$s_{\text{phon}} \ll s_n, \quad s \approx s_n. \tag{9}$$

The following considerations depend essentially on this assumption which will prove to be correct for not too low temperatures.

The assumption of reversibility of the superfluid flow implies that the flow across the capillary should be slower than the so-called critical velocity marking the beginning of dissipation.

Let us now compress the cell adiabatically by applying the pressure dP on the semi-permeable piston until the osmotic pressure difference balances the external pressure $dP_n = dP$. The condition of equilibrium with respect to the superfluid mass transfer is the equality of the chemical potentials:[20]

$$\mu(T, P) = \mu(T+dT, P+dP).$$

Expanding the right-hand side and noting that $(\partial\mu/\partial T)_P = -s \approx -s_n$, $(\partial\mu/\partial P)_T = v = 1/\rho$ (v is the specific volume), one has

$$(\partial P_n/\partial T)_\rho = \rho s_n. \tag{10}$$

The fact that the temperature difference is associated with an osmotic pressure is referred to as the thermo-mechanical effect or fountain effect. Relation (10) was first derived by H. London.[5,21]

While the derivation of (9) was worded in the terminology of the two-component system, an interpretation from the two-phase point of view is equally possible: Eq. (10) is then the Clausius-Clapeyron relation. Clearly, it should contain the entropy *difference* of the two phases, i.e., s_n rather than s.

The experimental verification of (10) is rendered difficult because of the necessity of maintaining reversible conditions. This seems to have been achieved to a sufficient degree by Kapitza[11] who found the relation (10) to be valid within experimental error.

Another important relation can be derived in connection with the osmotic cell. In the course of the isopycnic compression the normal component of the fluid and the entropy associated with it are "trapped" within the cell. Denoting the volume of the cell by V, one is thus led to the following relations:

$$d(\rho_n V) = \rho_n dV + V d\rho_n = 0$$

and

$$d(\rho s_n V) = \rho(s_n dV + V ds_n) = 0.$$

Hence

$$d\rho_n/\rho_n = ds_n/s_n, \tag{11}$$

or, on integration,

$$\rho_n/\rho = s_n/s_0, \tag{12}$$

[19] The small thermal expansion is neglected throughout this section.

[20] This follows from the principle of virtual work. It is somewhat unusual that the equality of chemical potentials is obtained for states at different temperatures. In the case of the true osmotic cell, the virtual process is the mass transfer of the solvent which is, of course, associated with transfer of energy (entropy). Hence, the equality of temperatures follows simultaneously with the equality of the chemical potentials. In the present case the superfluid flow carries no entropy or energy in excess over the zero-point value, thus the equality of the chemical potential does not imply the equality of temperatures.

[21] London's interpretation of Eq. (10) differs, however, in several respects from that of the present author. London did not distinguish between the pressures P_n and P, and assumed that the thermo-mechanical pressure depends essentially on the proximity of solid walls. The present interpretation is essential for the theory of second sound, to be discussed in the next section.

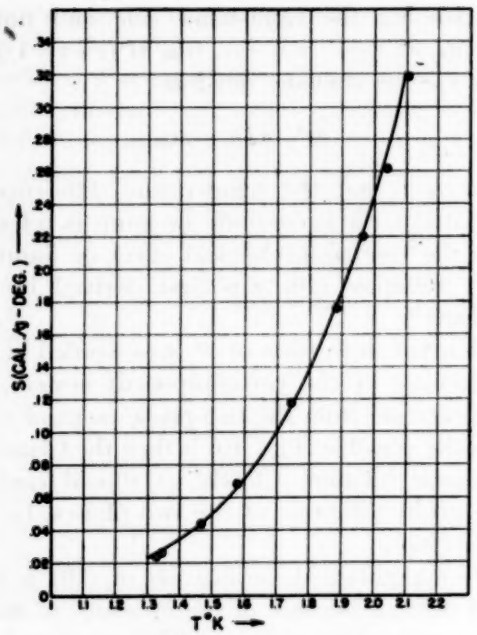


FIG. 3. Approximation of Kapitza's entropy measurements[11] (circles) with formula (13) (curve).

where $s_0 = s_n(T_0)$ is the entropy at the λ-point. This relation was obtained also by F. London.[22]

The function $s_n(T)$ plays the role of an equation of state for helium II and is of great importance for correlating its different properties. An analytic expression for this function will prove useful, we will use the expression

$$s_n = s_0(T/T_0)^r \qquad (13)$$

with $s_0 = 0.405$ cal./g-deg., $r = 5.5$. It is seen from Fig. 3 that (13) represents to a good approximation the entropy measurements of Kapitza. The entropy values following from the Leyden specific heat measurements seem to be a few percent lower but have essentially the same temperature dependence. Equation (13) has certainly only a preliminary character and should be corrected when new measurements become available.

Equations (12) and (13) confirm the expression (8) for ρ_n/ρ with a more precise value for the exponent r.

We are now in a position to discuss the validity of the condition (9). The contribution of the phonons to the entropy can be tentatively represented by the Debye formula, hence $s_{\text{phon}} \sim (T/\theta)^3$. One sees that (9) would certainly fail at sufficiently low temperatures provided that (13) can be extrapolated to this region. The temperature for which $s_{\text{phon}} \approx s_n$ could be determined if the Debye temperature θ were known. The sound velocity (~240 m/sec.) leads in the usual way to $\theta \approx 18°$. Picard and Simon[23] have actually found an approximate T^3 law for the specific heat in the temperature range 0.2°–1.0° but their measurements correspond to $\theta \approx 11°$.[24] This value of θ may be too low, the specific heat too high; actually Keesom and Westmijze (p. 220 of reference 5) report a lower value for the specific heat. While new measurements of the specific heat at low temperatures are desirable, it seems that the two contributions to the entropy are of the same order of magnitude for temperatures somewhere between 0.6° and 1°K. In the extreme low temperature region one has $s_n \ll s_{\text{phon}}$; the effect of the normal component can be neglected and the liquid becomes essentially homogeneous. The main characteristic of this region is its superfluidity no longer complicated by the viscosity of the normal component and the thermo-mechanical effect.

IV. MACROSCOPIC HYDRODYNAMICS IN HELIUM II

We will now derive the general macroscopic (coarse-grained) hydrodynamical equations of helium II in the range of temperatures for which condition (9) is satisfied. Only the so-called first approximation will be considered and dissipation effects will be neglected.

The state of motion of the liquid will be described by two vector fields ξ_n, ξ_s denoting the displacement of every elementary mass of the normal and superfluid liquids, respectively, from a standard state of constant densities ρ_{n0}, ρ_{s0}. The previously introduced velocities are $\mathbf{v}_n = d\xi_n/dt = \dot{\xi}_n$, $\mathbf{v}_s = d\xi_s/dt = \dot{\xi}_s$. Since we are considering processes in the first approximation, the densities ρ_n, ρ_s can change only by means of the *flow* of the components and non-adiabatic

[22] F. London, Rev. Mod. Phys. **17**, 310 (1945).

[23] G. L. Picard and F. Simon, Abstr. of papers comm. to the Roy. Soc. London S21 (3 April, 1939); compare R. Bleaney and F. Simon, Trans. Faraday Soc. **35**, 1205 (1939).

[24] Picard and Simon claim that their measurements should be described by $\theta = 15.5°$. This value is $3^{\frac{1}{3}}$ times too high since they used the Debye formula valid for solids, assuming two transverse waves for every longitudinal one. The transverse waves are absent in helium.

transitions from one component to the other should be ignored. Hence separate equations of continuity are valid for the two components:

$$\Delta\rho_n/\rho_n+\nabla\cdot\xi_n=0, \tag{14a}$$

$$\Delta\rho_s/\rho_s+\nabla\cdot\xi_s=0, \tag{14b}$$

where $\Delta\rho_n=\rho_n-\rho_{n0}$ is assumed to be small compared to ρ_{n0} and similarly $\Delta\rho_s\ll\rho_{s0}$. If Eqs. (14) are differentiated with respect to the time, one gets the equations of continuity in the conventional form.

The Lagrangian of the system is an integral over the volume

$$\mathfrak{L}=\int L\,dV, \tag{15}$$

where L is the Lagrangian density; i.e., the difference of the kinetic and potential energy densities. The kinetic energy density is $\frac{1}{2}\rho_n\dot{\xi}_n^2+\frac{1}{2}\rho_s\dot{\xi}_s^2$. A potential energy arises because of fluctuations of the specific volume and entropy from their values in the standard state. According to well-known results of thermodynamics[25] this potential energy is equal to the minimum work required to bring forth the fluctuations Δv, Δs from the standard state. Its density is given by

$$\tfrac{1}{2}\rho[(\partial^2u/\partial v^2)(\Delta v)^2+2(\partial^2u/\partial v\partial s)(\Delta v)(\Delta s)+(\partial^2u/\partial s^2)(\Delta s)^2], \tag{16}$$

where u is the energy per unit mass.

New variables will now be introduced in terms of which both kinetic and potential energy will take a simple form:

$$\xi_1=(\rho_n\xi_n+\rho_s\xi_s)/\rho, \tag{17a}$$

$$\xi_2=\xi_n-\xi=(\xi_n-\xi_s)\rho_s/\rho. \tag{17b}$$

A simple calculation using (14), (17), and (11) leads to the equations of continuity in terms of ξ_1 and ξ_2:

$$\Delta\rho/\rho+\nabla\cdot\xi_1=0, \tag{18a}$$

$$\Delta s_n/s_n+\nabla\cdot\xi_2=0. \tag{18b}$$

The physical meaning of the new variables is apparent from (17) and (18). The "center of mass" coordinate ξ_1 corresponds to the net flow of matter at constant entropy and the "relative coordinate" ξ_2 to the flow of entropy at constant density. The possibility of the latter process distinguishes helium II from ordinary homogeneous liquids for which one would have $\xi_2=0$.

The three terms of the potential energy will now be rewritten by using simple thermodynamic transformations. One has

$$(\partial^2u/\partial v^2)_s(\Delta v)^2=(\partial P/\partial\rho)_s[(\Delta\rho)/\rho]^2, \tag{19a}$$

$$(\partial^2u/\partial v\partial s)(\Delta v)(\Delta s)=\alpha Ts/c_v[(\Delta\rho)/\rho][(\Delta s)/s], \tag{19b}$$

$$(\partial^2u/\partial s^2)(\Delta s)^2=-(\partial T/\partial(1/s))_\rho[(\Delta s)/s]^2, \tag{19c}$$

where α is the coefficient of thermal expansion, c_v the specific heat at constant volume. Remembering that $s\approx s_n$ and by using (18), the potential energy is expressed in terms of $\nabla\cdot\xi_1$ and $\nabla\cdot\xi_2$. Finally, expressing the kinetic energy in terms of ξ_1, ξ_2, one obtains for the Lagrangian:

$$\mathfrak{L}=\int\tfrac{1}{2}\rho[\dot{\xi}_1^2+(\rho_n/\rho_s)\dot{\xi}_2^2-(\partial P/\partial\rho)_s(\nabla\cdot\xi_1)^2+(\partial T/\partial(1/s_n))(\nabla\cdot\xi_2)^2-\alpha Ts_n/c_v(\partial P/\partial\rho)_T(\nabla\cdot\xi_1)(\nabla\cdot\xi_2)]dV. \tag{20}$$

An alternative form for (19b) and hence for $\mathfrak{L}$ is obtained by noting that from (10) and (11) one gets

$$-(\partial T/\partial(1/s_n))_\rho=(\rho_n/\rho)(\partial P_n/\partial\rho_n)_\rho. \tag{21}$$

The equations of motion corresponding to (20) are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\xi}_1}\right)+\nabla\left(\frac{\partial L}{\partial(\nabla\cdot\xi_1)}\right)-\frac{\partial L}{\partial\xi_1}=0 \tag{22}$$

and similar equations in ξ_2, or explicitly:

$$d^2\xi_1/dt^2-(\partial P/\partial\rho)_s\nabla(\nabla\cdot\xi_1)=((\alpha Ts_n)/c_v)(\partial P/\partial\rho)_T\nabla^2\xi_2, \tag{23a}$$

$$d^2\xi_2/dt^2+(\partial T/\partial(1/s_n))_\rho\nabla(\nabla\cdot\xi_2)=((\alpha Ts_n)/c_v)(\partial P/\partial\rho)_T(\rho_s/\rho_n)\nabla^2\xi_1. \tag{23b}$$

A numerical calculation based on the observed values of the quantities involved shows that the right-hand sides of (23) are very small. These terms representing a coupling between the ξ_1 and ξ_2 motions will hence be neglected. Equations (23) are then simply the Eulerian equations in first approximation for the two flows, respec-

[25] Compare L. Landau and E. Lifshitz, *Statistical Physics* (Oxford University Press, New York, 1938), p. 100.

tively. In fact, by using (18) and (21) one gets:

$$d\dot{\xi}_1/dt=\partial\dot{\xi}_1/\partial t+(\dot{\xi}_1\cdot\nabla)\dot{\xi}_1=-(\nabla P)/\rho, \quad (24a)$$

$$d\dot{\xi}_2/dt=\partial\dot{\xi}_2/\partial t+(\dot{\xi}_2\cdot\nabla)\dot{\xi}_2 =-(\nabla P_n)(\rho_s/\rho\rho_n). \quad (24b)$$

By making use of the vector relation $\nabla(\nabla\cdot\xi)=\nabla^2\xi+\nabla\times(\nabla\times\xi)$, taking the divergence of the Eqs. (23) justified and replacing the total time derivatives by partials (justified for small velocities), one obtains two wave equations:

$$\partial^2(\nabla\cdot\xi_1)/\partial t^2-c_1^2\nabla^2(\nabla\cdot\xi_1)=0 \quad (25a)$$

$$\partial^2(\nabla\cdot\xi_2)/\partial t^2-c_2^2\nabla^2(\nabla\cdot\xi_2)=0 \quad (25b)$$

The value of the first wave velocity is

$$c_1=[(\partial P/\partial\rho)_s]^{\frac{1}{2}}, \quad (26)$$

and, by (21), the second wave velocity can be given in either of two alternative forms:

$$c_2=[(\partial P_n/\partial\rho_n)\rho_s/\rho]^{\frac{1}{2}}, \quad (27a)$$

$$c_2=[-(\partial T/\partial(1/s_n))_\rho\rho_s/\rho_n]^{\frac{1}{2}}. \quad (27b)$$

Finally, by using (12) and the empirical relation (13), one obtains c_2 as a function of temperature:

$$c_2=26[(T/T_0)(1-(T/T_0)^{5.5})]^{\frac{1}{2}} \text{ m/sec}. \quad (28)$$

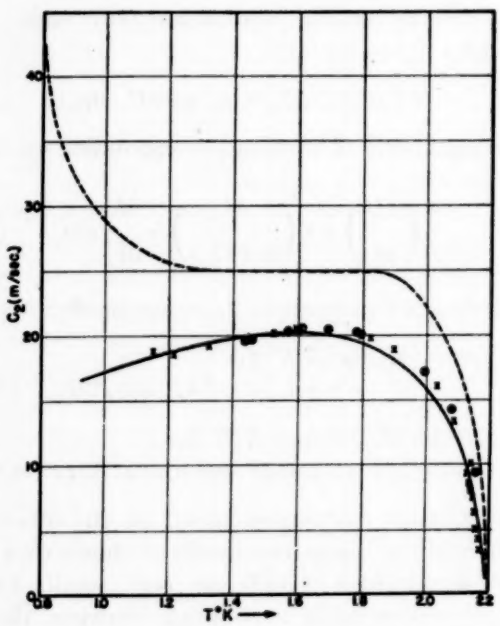


FIG. 4. Velocity of second sound. xxx-measurements of Peshkov.[8] ooo-measurements of C. T. Lane and collaborators.[10] —— formula (28) of present paper. – – – calculated by Lifshitz from Landau's theory.[9]

The physical meaning of the two waves is apparent from Eqs. (18), (8), (11). It is seen that $\nabla\cdot\xi_1$ is equal to the fluctuation of the density ρ and hence also of the pressure P. The corresponding wave has the properties of ordinary wound waves. The ξ_2 waves consist of fluctuations in s_n, ρ_n, P_n, and T. The two waves can be conveniently referred to as pressure and temperature waves,[4] or simply first and second sound.[7] The terminology of temperature waves is suggestive of the simplest way of generating these waves by periodic heating, which has actually been carried out by Peshkov[8] and by Lane and his collaborators.[18]

In the case of heat transfer between the liquid and its surroundings, the continuity equation (18b) has to be supplemented by "sources" for the normal component. If the heating occurs across a solid wall, this source can be conveniently accounted for by the boundary conditions for the perpendicular components of $\dot{\xi}_1$ and $\dot{\xi}_2$:

$$\dot{\xi}_{1\perp}=0, \quad \dot{\xi}_{2\perp}=W/Ts_n\rho, \quad (29)$$

where W is the amount of heat absorbed by the liquid across a unit area per unit time.

The very existence of the temperature waves bears out the fundamental assumptions of the theory.[26] A more quantitative check of the theories is obtained by comparison of the theoretical expressions of $c_2(T)$ with the observed values, as shown in Fig. 4. The agreement is seen to be very satisfactory. It should be noted that formula (28) contains no adjustable constants.

The figure also shows the function $c_2(T)$ computed by Lifshitz[9] on the basis of Landau's

[26] Recently E. G. Richardson, Nature **158**, 296 (1946) expressed the view that the temperature waves might be explained within the ordinary theory of heat conduction simply by a large value of the coefficient of heat conductivity χ, and without assuming any wave equation for the temperature. Although the parabolic equation of heat conductivity actually possesses wave solutions, these waves are strongly damped (independently of the value of χ) and show a strong dispersion, their velocity of propagation being proportional to $\omega^{\frac{1}{2}}$ (ω is the frequency). The criterion of dispersion for distinguishing between the wave solutions of a parabolic and a hyperbolic differential equation has been pointed out previously.[4] The experiments[8] decide the question without ambiguity in favor of the hyperbolic equation. It may be also mentioned that the true heat conductivity in helium II is proportional to the viscosity and is of the same order of magnitude in helium II as in helium I. Its presence is completely masked by the presence of the internal convection.

theory. The reason for the discrepancy will be discussed in Section VI.[27]

The agreement between (28) and the experimental points is not complete. This is believed to be due mainly to inaccuracy of the empirical relation (13), rather than to any inadequacy of (27). It is true that the neglected dissipation effects will give an absorption and dispersion of the temperature waves and hence a modification of (27), but these effects seem to be small since Peshkov failed to find any dispersion from 30 to 10,000 c.p.s.

The discussion of the dissipation effects is beyond the scope of this theory, but the various effects should now briefly be listed.

In addition to dissipation caused by viscosity and the true heat-conductivity, there are effects characteristic of helium II. The first of these is due to non-adiabatic transitions between the two components which lead to deviations from (18b).

A second effect is due to the Debye phonons which should give rise to an absorption and dispersion of the relaxation type. This can be expected to be of importance for the second sound at low temperature ($T<1°$K) and should eventually (at sufficiently low temperatures) lead to a complete attenuation of the temperature wave within one wave-length. In case the "molecules" of postulate (*c*) should have internal degrees of freedom (vortex model), another relaxation phenomenon should be expected. Considerable dissipation arises if the flow velocity exceeds the so-called critical value. This seems to be of a turbulent character.[28]

A quantitative refinement of this theory should take into account the small coupling between the ξ_1 and ξ_2 motions and also the fact that ρ_n/ρ and s_n/s_0 are functions of the pressure and not only of the temperature.

This section will be closed by a brief outline of the qualitative discussion of the experiments on heat conductivity and capillary flow.

Let us first consider the heat conductivity. Two heat reservoirs of temperatures T_1, T_2 are connected by a capillary completely filled with helium II precluding any net flow of the liquid. Hence $\dot{\xi}_1=0$. The temperature difference $\Delta T=T_1-T_2$ gives rise to a difference of osmotic pressure ΔP_n and thus to a circulation in the capillary ($\dot{\xi}_2\neq 0$), whereby the normal component flows towards the cold reservoir. The current is closed by a transition of the superfluid component into the normal at the warm reservoir while absorbing the heat given by (29). The reverse process takes place at the cold reservoir. The efficiency of this heat transport is about $T/\Delta T$ times larger than ordinary convection. This factor may reach very high values of the order of 1000. In the steady state the flow is limited by dissipation effects neglected in the above discussion. If the dissipation were due to the viscosity of the normal component alone, the heat current should be proportional to ∇P_n and hence to ∇T. Assuming *tentatively* Poiseuille's law for the internal convection, one gets for the heat flow

$$W=(\pi r^4 s^2\rho^2\rho_s)/(8\eta\rho_n)T\nabla T, \qquad (31)$$

where r is the radius of the capillary. Actually the heat current increases rather like $(\Delta T)^4$, (reference 5), which makes it likely that the dissipation is partly of a turbulent character. This is supported by a series of experiments carried out by Kapitza.[28]

The situation is even more complicated in the case of capillary flow. In the limiting case of very thin capillaries only the superfluid component can flow under the influence of a pressure head, say caused by gravity. Thus a temperature difference is set up between the two ends of the capillary, the outflowing liquid being colder than that staying behind.[29] The flow should thus be stopped by the osmotic pressure difference which arises. Actually the increased temperature will lead to vaporization of the liquid with a corresponding cooling and transition of the normal component to the superfluid with a continuation of the flow. In somewhat wider capillaries the

[27] It may be noted that (27a) was first given by the author.[3] Also (28) was then given except for an undetermined constant of the order of unity. A formula similar to (27b) was derived by Landau,[7] but he obtained $s=s_n+s_{phon}$ in place of s_n. Expression (28) was first given by F. London in a review paper to be published in the Proc. Phys. Soc. The author is indebted to Dr. London for communicating his paper before publication.

[28] P. Kapitza, J. Phys. U.S.S.R. **4**, 181 (1941).

[29] This was one of the first conclusions of the theory which was readily verified by experiment. J. G. Daunt and K. Mendelssohn, Nature **143**, 719 (1939). The outflowing liquid is never at absolute zero as it still contains the phonon entropy, and at least a small fraction of the normal component.

normal flow velocity also becomes appreciable and both flows are complicated by turbulence. Consequently, the thermo-mechanical pressure will be smaller than that expected from Eq. (10), which was derived under the assumption of strictly semi-permeable capillaries. In summing up, it is clear that the complexity of the results obtained with all but the narrowest capillaries is not surprising and the disentanglement of the various effects from these measurements is hardly feasible.

It would seem advisable to study the dissipative effects by measuring the absorption of first and second sound whereby the complications of turbulence are avoided. On the other hand, the study of turbulence could best be carried out at low temperatures, where the concentration of the normal component is negligible and the flow is no longer complicated by the thermo-mechanical effect and the viscosity of the normal component.

V. OUTLINE OF A MICROSCOPIC THEORY OF LIQUID HELIUM

The conclusions drawn from the postulates (*a*)–(*d*) are in good agreement with experiment; practically all the peculiar kinetic effects finding either a quantitative or at least a qualitative interpretation. Thus the task of a microscopic theory is reduced to providing a quantum-mechanical foundation for these postulates. There is at present no question of deriving the postulates from first principles, and the following discussion has only an intuitive character.

No special discussion is required for postulate (*a*) since the presence of van der Waals attractive and repulsive forces always assures its validity. In contrast to the universal nature of (*a*), the validity of postulate (*b*) is most exceptional. Systems satisfying both postulates have been called quantum liquids in Section II.

It would be desirable to find criteria for the interatomic forces which assure that the macroscopic system built of these atoms will be a quantum liquid. Although this does not seem possible at present, one can point out a peculiarity of the repulsive forces in liquid helium which seems to be at the root of this question. In an ordinary (classical) system both attractive and repulsive forces between the atoms are of the van der Waals type. In helium, however, an additional repulsive force appears which originates in the quantum-mechanical zero-point energy. It blows up the volume of the liquid to about three times the value one would expect from the atomic diameter. The importance of the zero-point energy was first recognized by F. Simon,[30] and the idea has been developed by F. London.[31] For further details we refer to these papers.

The inference from these considerations is that the absence of microscopic rigidity as postulated in (*b*) might be due to the quantum-mechanical nature of the repulsive forces. Whereas in an ordinary liquid the potential barriers opposing a shearing motion (isopycnic rearrangements of the molecules) break down under the impact of thermal agitation, in helium this happens as a result of the zero-point motion. No proof of this statement has been given so far but the following remark may prove relevant.

Consider a particle in a *cubic* box of volume V. The box is slightly deformed into a rectangular parallelepiped of volume $V+\Delta V$. The zero point energy appears here as the lowest eigenvalue E_0 of the particle. It can be easily shown that, neglecting quadratic terms in the deformation, one has

$$\Delta E_0/E_0 = -\Delta V/V. \tag{30}$$

Hence, for a shearing deformation ($\Delta V=0$), the energy is in first approximation unaffected. More generally, (30) holds for an energy level of the cubic enclosure if, and only if, the corresponding ψ function is invariant under the cubic rotation $x\rightarrow y\rightarrow z$.

Although this case is altogether too simple to be applicable to the many-body problem of liquid helium, it is hard to suppress the feeling that this selective behavior of the zero-point energy with respect to shear and compression might be of more general validity.

We turn now to the discussion of the postulates (*c*) and (*d*) from the point of the Bose-Einstein model.

As first suggested by London,[2] individual helium atoms might exist in liquid helium in

[30] F. Simon, Nature **133**, 529 (1934).
[31] F. London, Proc. Roy. Soc. **A153**, 576 (1936); J. Phys. Chem. **43**, 49 (1939); compare also T. Nagamiya, Proc. Phys. Math. Soc. Japan (3), **22**, 492 (1940).

excited translational states associated with a definite energy and momentum in much the same way as the electrons in metals according to Bloch's theory. The Bose-Einstein statistics implies a "condensation in momentum space," hence postulate (*d*) appears as a theorem on the basis of this model.

The incorporation of London's hypothesis in the present theory considers only the shear modes of motion, and otherwise the system has the liquid characteristics as discussed above. Such a system can be conveniently called a *Bose-Einstein* liquid.[4] It is assumed that its ψ-function in a state of weak excitation can be written as

$$\Psi=\psi_0(1, 2, \cdots, N)\psi_{\text{phon}}\psi_{\text{gas}},$$

$$\psi_{\text{phon}}=\prod_i \psi_i(u_i), \tag{31}$$

$$\psi_{\text{gas}}=\sum_P \prod_{j=1}^{N_{\text{exc}}} \exp[i\mathbf{k}_j\cdot\mathbf{r}_j].$$

The function $\psi_0(1, 2, \cdots N)$ is essentially the eigenfunction of the lowest state. *It depends on the coordinates of all the atoms.* The u_1 are the longitudinal normal modes of the system also involving the coordinates of all the atoms. In contrast to this situation ψ_{gas} depends only on the coordinates of the excited atoms. (The assumption of weak excitation implies $N_{\text{exc}}\ll N$.) The symbol $\sum_P$ represents the summation over all the permutations of the particles.

The form (41) for the ψ-function has not been derived from first principles, and it is hard to see how that could be done at present. All we intend to point out is that it is conveivable to think of a system of strongly interacting particles where the ψ-function corresponding to certain modes of motion can be factorized nevertheless. This is an important point in view of the numerous objections which had been brought up against the Bose-Einstein theory.

First, it should be pointed out that Keesom[5] was justified in rejecting the interpretation of the viscosity measurements given in reference 4. This objection concerned, however, only an erroneous application of the theory, not the theory itself. After removal of this error (see this paper, Section III and the discussion of the free path given below) the viscosity measurements of Keesom and MacWood give a particularly striking support of the theory. Apart from this instance, most of the criticism seems to stem from some misunderstanding the main source of which is in the failure to appreciate the difference between a real Bose-Einstein *gas* and a Bose-Einstein *liquid*. Several attempts have been made to replace the *ideal* by a *real* Bose-Einstein gas by introducing a self-consistent potential. This method fails to account for the correlations between the states of the individual particles which are all important in the case of a liquid. They are taken into account to a large extent by our postulates: the gaseous nature of the system is restricted to the shear modes of motion (isopycnic rearrangements), its compressibility is that of a liquid. This remark disposes of the numerous attempts which tended to prove or disprove the theory by means of the "critical opalescence" arising from the large density fluctuations in a condensed *Bose-Einstein gas*.[32] In a *Bose-Einstein liquid* the density fluctuations are determined by the liquid-type compressibility (observed, for example, in the first sound), which shows no significant anomalies around the λ-point.

The fact that the Bose-Einstein liquid allows one to take into account the correlations between the excited atomic states to a satisfactory degree is due to the artifice of formulating these correlations macroscopically rather than in quantum-mechanical terms. In other words, the difficulties of the quantum-mechanical many-body problem are not *solved*, but *by-passed* in the present theory.

Another source of misunderstanding is that in the Bose-Einstein liquid one has to distinguish "collisions" of the excited atoms from "interactions" of the atoms in general.

All atoms are strongly interacting through the mechanism of the zero-point energy precluding large fluctuations of density. In other words, the factor ψ_0 in (31) is not separable in the coordinates of the particles. In contrast to this interaction, one speaks of a "collision" if two or more atoms in definite translational states interact to make a transition into different states. In such processes translational energy

[32] L. Goldstein, Phys. Rev. **57**, 241, 457 (1940); L. I. Schiff, Phys. Rev. **57**, 844 (1940); V. L. Ginsburg, J. Phys. U.S.R.R. **7**, 305 (1943).

and momentum are conserved. Obviously double collisions between excited and non-excited atoms are excluded by these conservation principles which assure an apparent independence of the two components. This is the basis for the assumption $\rho_n l = \text{constant}$, which was advanced in Section III regarding the mean free path. The mechanism for the establishment of thermal equilibrium is given by multiple collisions. These tend to maintain the value appropriate to the temperature.

Finally, we conclude that the Bose-Einstein theory furnishes a self-consistent although hypothetical molecular model for the postulates (*a*)–(*d*). It is, of course, of importance to know whether there is another molecular model with the same properties. This will be discussed in the next section.

VI. DISCUSSION OF LANDAU'S THEORY

No attempt will be made here to analyze the logical connection between Landau's microscopic quantum hydrodynamics on the one hand and his macroscopic relations on the other. These relations will be considered rather from the point of view of our quasi-thermodynamic theory and the question to be answered is: What changes and additions have to be applied to the postulates (*a*)–(*d*) in order to obtain Landau's results?

Postulates (*a*) and (*b*) are made tacitly by Landau. In particular, the validity of (*b*) is essential if the Hamiltonian of the system should depend on the density alone as is implied by formula (1.10) of reference 6.

In Landau's theory, the excited atoms of the Bose-Einstein theory are replaced by "rotons." A divergence from the present theory arises in connection with postulate (*d*). There is agreement inasmuch as helium II has to be considered as a mixture of a normal and a superfluid component, but Landau defines the normal component as a "gas of phonons and rotons," leaving the superfluid component with rigorously vanishing entropy.

It is seen that the theories agree sufficiently closely to account for the similarity of the results. The difference in the role attributed to the phonons leads, however, to several observable discrepancies.

In the first place, Landau fails to get the condition of validity (9) for the thermo-mechanical effect. Accordingly Kapitza[11] suggested a cooling method which would *a priori* permit us to approach infinitely near to the absolute zero. The investigation of this effect at low temperatures should allow us to test the importance of condition (9).

A second point concerns the expression (27b) for the second sound velocity. The function $c_2(T)$ depends very sensitively on the manner ρ_n/s_n depend on temperature. Equation (12) implies generally $c_2 \sim T^1$ for $T \to 0$. Landau, in contrast, concludes that so far as phonons are concerned, $\rho_n \sim u_n \sim T^4$ where u_n is the phonon energy, while the corresponding entropy is $s_n \sim T^3$. This is a violation of (12) and leads in the limits of low temperatures to $c_2 \to c_1/3^{\frac{1}{2}} \approx 130$ m/sec. Thus the discrepancy between Landau's theory and the experiments (compare Fig. 4) originates in the violation of Eq. (12).[33]

Landau's discussion of phonons can be rejected already from the point of view of the quasi-thermodynamic theory. The question of rotons is more difficult since the macroscopic theory remains unchanged if the "elementary excitations" of postulate (*c*) are interpreted in terms of "rotons" rather than excited atoms. The merits of this interpretation have to be judged from more specific results. We find two results of this kind in Landau's theory. First, the exponential temperature dependence of ρ_n and s_n, and second, the high value of about 100 m/sec.

[33] After this manuscript had been completed, Landau published a note (J. Phys. U.S.S.R. **11**, 91 (1947) where, in an attempt to account for this discrepancy, he suggested a new expression for the energy spectrum of helium. This expression was obtained by adjustment of three arbitrary constants. Apart from the unsatisfactory nature of this procedure, it tends to modify the theory in the wrong direction. Originally Landau failed to notice that every vortex element can be associated with a definite mass contained in the volume in which the vorticity is different from zero. It is a consequence of the hydrodynamic equations that this mass is an integral of the motion. In contrast to this situation, phonons are associated with the liquid as a whole.

Only elementary excitations associated with definite masses can make it understandable that the liquid breaks up into two components as required in postulate (*d*). Landau introduced such a mass by his device of the rotating vessel for rotons *and* phonons *indiscriminately*. This argument is not convincing as it tends to obtain information on a kinetic coefficient (viscosity) from equilibrium considerations. Thus Landau had ignored an important difference between phonons and rotons. In his latest paper this difference is even more blurred.

for the critical velocity. Both results are in conflict with experiment; they follow directly from Landau's assumption of a "gap" in the vortex spectrum. If this assumption is dropped, new difficulties arise, the discussion of which is beyond the scope of this paper.

Summing up, it may be said that the vortex model is in a rather unsatisfactory state. In some respects, however, the situation is not unlike that encountered in the Bose-Einstein theory. A rigorous proof or disproof of the theory seems equally difficult. Under such circumstances, it is of interest to see whether the question could be decided experimentally. Even this approach is rather difficult, since one has to consider *second-order* hydrodynamic effects which are at present little understood. A distinguishing feature of the vortices compared with the excited atoms of the Bose-Einstein theory is the existence of an internal rotational energy which should manifest itself in an absorption and dispersion of the relaxation type for first and particularly for second sound.

Interesting crucial experiments would be possible if the H^3 isotope obeying Fermi statistics would be sufficiently enriched to make flow experiments possible. Such experiments should show conclusively whether the Bose-Einstein statistics are of any fundamental importance for the phenomenon of superfluidity. If liquid He^3 should not be superfluid, this fact could be actually used for the enrichment of this isotope as has been recently suggested by J. Franck.[34]

VII. CONCLUSIONS

In the discussion of the properties of liquid helium, it is useful to distinguish three ranges of temperature in which the liquid shows essentially different characteristics. In addition to the well-known modifications helium I and II, separated by a sharp λ-point, there is a gradual transition between a low temperature and high temperature domain in helium II. The transition region is somewhat below 1°K, but its location needs further experimental study. In the low temperature region, the entropy is due mainly to the elastic phonons (elementary excitations of the compressional modes), whereas in the high temperature region the phonon contribution is negligible compared with that of the elementary excitations of shear modes (isopycnic rearrangements of the liquid). This contribution is responsible for the anomaly in the specific heat and has a "gaseous" nature. The main properties of the liquid are as follows:

Helium I

From the macroscopic point of view, this modification shows a "normal" hydrodynamic behavior; its coefficient of viscosity is, however, rather remarkable. In ordinary liquids the viscosity has a negative temperature coefficient because of the "microscopic rigidity." This type of viscosity can be called *dynamic viscosity*, to be distinguished from the *kinetic viscosity* which is observed in gases and which has a positive temperature coefficient. The qualitative understanding of the properties of liquids is greatly facilitated by the concept of *microscopic rigidity*. However, the difficulties of giving to this notion a quantitative formulation seem to be responsible in no small degree for the lack of a satisfactory kinetic theory of liquids. Hence it seems to be of interest that helium I has a kinetic and *no* dynamic viscosity revealing the complete absence of microscopic rigidity. The comparison of ordinary liquids with helium I should be useful for the disentanglement of the gas-like and solid-like properties of liquids.

Helium II. Low Temperature Regions ($0 < T \lesssim 1°K$)

The main characteristic of the liquid in this temperature range is its superfluidity which implies that the liquid can slip along solid walls. This is demonstrated in the spectacular creeping phenomena. Beyond a certain critical velocity, the superfluid flow may be associated with dissipation, apparently of a turbulent character.

Helium II. High Temperature Region ($1° \lesssim T < 2.19°K$)

In this "anomalous" region the liquid is a mixture of a normal component (like helium I) and a superfluid component (like the low temperature form of helium II). The main object of the present paper was to develop the macroscopic thermo-hydrodynamic properties of such a two-

[34] J. Franck, Phys. Rev. **70**, 261 (1946).

fluid system. The results are in excellent agreement with experiment. Minor discrepancies may well be due to small inaccuracies in the caloric measurements, although the theory will have to be refined to include small effects neglected at present, like dissipation and the dependence of ρ_n/ρ on pressure. The formalism developed seems to furnish valuable criteria for the consistency of measurements and might possibly be used to establish the thermodynamic temperature scale.

The principal problems for further experimental and theoretical research seem to be as follows:

(1) The derivation of the postulates of the quasi-thermodynamic theory from first principles. In particular, establishing the correct molecular model which would lead to these postulates. At present the choice seems to be between the Bose-Einstein liquid and possibly the vortex model (modified Landau theory). The experimental approach to this problem has been briefly discussed at the end of Section VI.

(2) *The clarification of the nature of the pure superfluid liquid (helium II near absolute zero).* This problem has not been discussed in the present paper. According to Landau, the superfluid state is characterized by the condition $\mathrm{curl}\mathbf{v}_s=0$. The question has been further discussed by F. London[18] and by Onsager.[35] So far, superfluidity (capillary flow and the creeping phenomenon) has been studied experimentally only above 1°K where the effects have been greatly complicated by the thermo-mechanical effect and the viscosity of the normal component. Experiments below 1°K where these effects practically vanish would be greatly desirable. Of particular interest is the investigation of the dissipation arising above the critical velocity. This seems to be of turbulent character; hence, a unique opportunity is offered for studying turbulence in the absence of viscosity (no boundary layer). On the other hand, this turbulence will be influenced by quantum effects.

The author wishes to extend his thanks to Dr. F. London for a long series of discussions extended over a period of years. They have led to the clarification of many of the questions considered in this paper. His thanks are also due Dr. Lars Onsager for stimulating discussions.

[35] Private communication.

PHYSICAL REVIEW VOLUME 72, NUMBER 9 NOVEMBER 1, 1947

Theory of the Ferroelectric Effect and Clamped Dielectric Constant of Rochelle Salt

W. P. MASON
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received July 25, 1947)

Using the crystal structure determination of rochelle salt made by Beevers and Hughes, a theory of the ferroelectric effect and clamped dielectric constant has been worked out using the displacement of the hydrogen nucleus in the 1–10 hydrogen bond as the ferroelectric dipole. This theory has only one disposable constant, β, the factor of proportionality between the polarization and the Lorentz internal field. Taking the oxygen separation found by x-rays, the dielectric constant for electrons and atoms found by experiment, and the number of molecules per cubic centimeter found from the x-ray cell determinations, the value of β becomes 4.07 which agrees well with the theoretical value $4\pi/3$ for an isotropic substance. This theory accounts for the clamped dielectric constant at low field strengths, which as shown by measurements presented here, has maximum values at −18°C and +24°C, the Curie temperatures for the free crystal. This shows that the anomaly must lie in the clamped dielectric constant alone and not in the interaction of the clamped dielectric constant and the piezoelectric effect. This model also accounts for recent measurements of W. A. Yager which show that the dipole dielectric constant is relaxed at a frequency of about 5×10^9 cycles.

I. INTRODUCTION

ALL measurements of the properties of rochelle salt indicate that the piezoelectric and elastic properties are normal and that all the anomalies reside in the "clamped" dielectric constant, i.e., the dielectric constant measured in the absence of strain. The piezoelectric stress has been found to be directly proportional to the

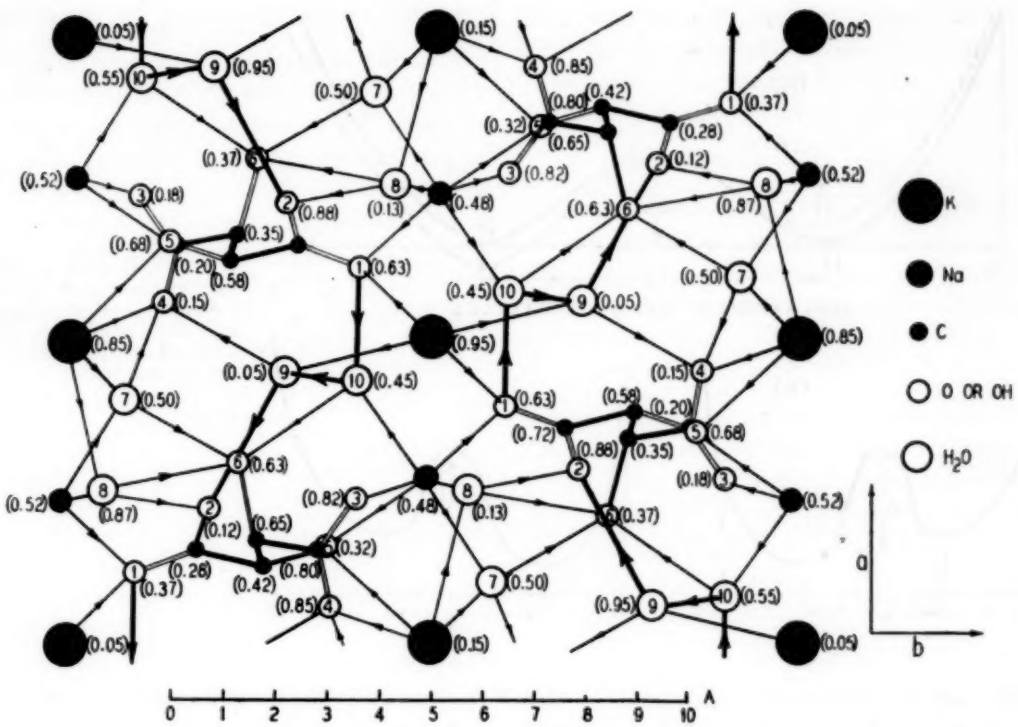


FIG. 1. Projection of (001) plane, of the structure of rochelle salt. The bonds involved in the 1–2–9–10 chains are drawn thick, so that the chains can be picked out.

Table of Coordinates

	$X=a$	$Y=b$	$Z=c$		$X=a$	$Y=b$	$Z=c$
2K on (*a*)	0.00	0.00	0.05	$4H_2O$ on (7)	0.40	0.08	0.50
2K on (*b*)	0.00	0.50	0.15	$4H_2O$ on (8)	0.25	0.05	0.87
4Na on	0.23	0.99	0.52	$4H_2O$ on (9)	0.44	0.30	0.05
4O on (1)	0.12	0.10	0.37	$4H_2O$ on (10)	0.42	0.40	0.45
4O on (2)	0.22	0.20	0.12	4C on	0.15	0.18	0.28
4O on (3)	0.23	0.40	0.82	4C on	0.12	0.28	0.42
4O on (4)	0.06	0.37	0.85	4C on	0.17	0.27	0.65
4OH on (5)	0.16	0.36	0.32	4C on	0.15	0.35	0.80
4OH on (6)	0.29	0.24	0.63				

dipole polarization and the large piezoelectric strain for rochelle salt is due to the large polarization for a given voltage, caused by the large clamped dielectric constant. Hence all theories of the ferroelectric effect in rochelle salt are primarily theories[1] of the clamped dielectric constant. The most recent and comprehensive of these theories is the interaction theory of Mueller,[2] which ascribes the spontaneous polarization between the Curie temperatures of −18°C and +24°C to an interaction between the clamped dielectric constant and the piezoelectric effect, which produces a unity electromechanical coupling between the mechanical displacement and the applied electric field when the clamped dielectric constant is high enough. This theory would indicate that the "monoclinally clamped" dielectric constant, i.e., the dielectric constant measured when the spontaneous polarization occurs but the frequency is so high that no piezoelectric motion can occur, would have Curie temperatures of −12°C and +18°C, i.e., well inside of the temperatures −18°C and +24°C. This is a prediction that can be tested experi-

[1] W. G. Cady, a review of the theories of rochelle salt is given in *Piezoelectricity* (McGraw-Hill Book Company, Inc., New York, 1946).

[2] H. Mueller, Phys. Rev. **57**, 829 (1940); *ibid.* Phys. Rev. **58**, 565 (1940).

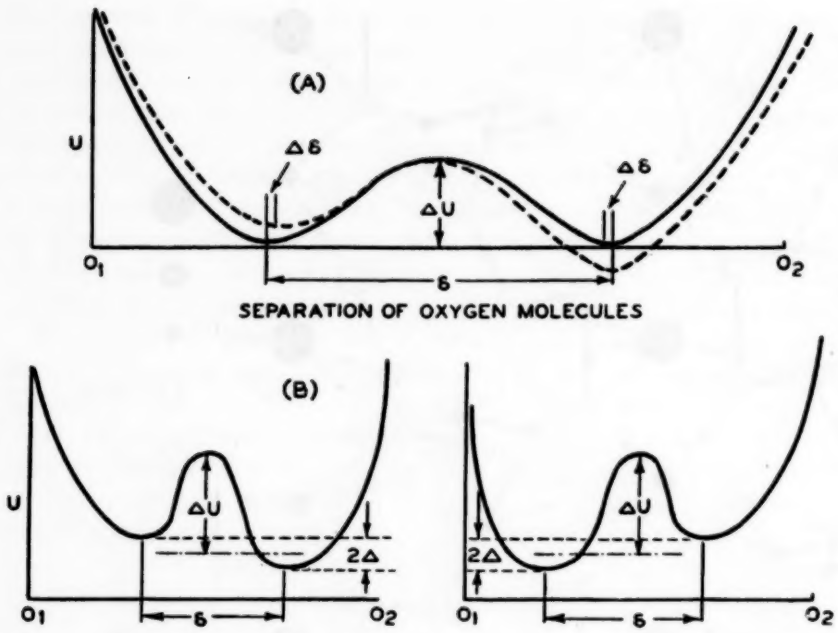


FIG. 2. Possible potential well distributions on 1–10 bond. (A) Potential well distribution if 1 and 10 are equivalent. (B) Potential well distributions if 1 and 10 are not equivalent.

mentally and measurements presented in this paper, made at 20 megacycles for large-sized crystals, do not show any displacement of the Curie temperatures from $-18°C$ and $+24°C$. Hence the anomoly cannot reside in the interaction between the piezoelectric effect and the clamped dielectric constant but must reside in the clamped dielectric constant alone.

Recently, a determination of the crystal structure of rochelle salt has been made by Beevers and Hughes,[3] which indicates that the dipole causing the ferroelectric effect is one associated with the motion of a hydrogen nucleus along a hydrogen bond. The structure of rochelle salt according to Beevers and Hughes is shown by Fig. 1. There are three possible hydrogen bonds in the structure, between the oxygen molecule 1 and the water molecule 10, between the water molecule 10 and the water molecule 9, and between the water molecule 9 and the oxygen 2. The distances between the successive molecules are: 1 to 10 is 2.59A;[4] 9 to 10 is 2.86a; and 9 to 2 is 3.02A. The bond with the shortest distance is 1 to 10 (2.59A) and is believed by Beevers and Hughes and also by Ubbelohde and Woodward,[5] to be the dipole responsible for the ferroelectric properties of rochelle salt. This bond lies nearly along the X axis (X coordinate $=2.35$A; Y coordinate $=0$: Z coordinate $=\pm 1.09$A) which is the ferroelectric axis of the crystal. It is the purpose of this paper to discuss a theory based on the action of this hydrogen bond, which leads to a ferroelectric effect having the right value of spontaneous polarization, two Curie temperatures, a good quantitative agreement with the measured values of the clamped dielectric constant and agrees well with recent measurements of W. A. Yager on the dielectric constant of rochelle salt at a frequency of 2.5×10^{10} cycles, which show that the dipole dielectric constant is relaxed at a frequency of about 5×10^{9} cycles. No phenomenological constants are required and only one disposable constant is required, namely, β which is the factor of proportionality between the polarization and the Lorentz internal field.

[3] C. A. Beevers and W. Hughes, Proc. Roy. Soc. **A177**, 251 (1941).

[4] The distance 1–10 calculated from Fig. 1 is 2.64A; but Ubbelohde and Woodward report that the most probable distance is 2.59 ± 0.05A.

[5] A. R. Ubbelohde and I. Woodward, Proc. Roy. Soc., **A185**, 448, 465 (1946). They suggest that the lower Curie point is caused by the disappearance of the two potential wells in the hydrogen bond. This cannot be true, for as shown in this paper such a disappearance would eliminate the temperature variable dipole below the lower Curie point and the measured dielectric constant should be independent of temperature. This does not agree with experiment.

II. FERROELECTRIC EFFECT IN ROCHELLE SALT

We suppose that in the undisturbed state, the hydrogen nucleus in the hydrogen bond 1 to 10, which lies at an angle of about 25° with the X axis of the crystal, is in the potential field shown by Fig. 2A. There are two minima, a potential barrier ΔU between them and a distance δ separating the two minima. Now according to elementary kinetic theory, the probability of a nucleus in one potential well jumping to the other potential well, per unit of time, is

$$\alpha=\Gamma e^{-\Delta U/kT}, \tag{1}$$

where Γ is a constant. (In Eyring's reaction rate theory $\Gamma=kT/h$.) Let α_+ be the probability of a particle jumping in the positive direction per unit of time and α_- be the probability of its jumping in the negative direction, and let N_+ be the population in the wells along the positive direction and N_- the population in the wells in the negative direction. Then the dipole polarization is equal to

$$(N_+-N_-)\mu=P_d, \tag{2}$$

where μ is the dipole moment along the X axis of the bond and neighboring atoms. Since the bond is at an angle of 25° from the X axis $\mu=\mu_A\cos 25°$ where μ_A is the absolute value of the dipole along the bond. We have also

$$N_++N_-=N, \tag{3}$$

the total number of dipoles per unit volume.

The rate at which the dipole polarization changes with time is

$$dP_d/dt=[N_-\alpha_+-N_+\alpha_-]\mu, \tag{4}$$

or it is the number of molecules in the wells along the negative direction times their probability of jumping to the positive wells minus the reverse reaction, all multiplied by the dipole moment μ. Now suppose that we put on a field E. This will change the potential wells as shown by the dotted line of Fig. 2. There is a change $\Delta\delta$, in the position of the minima and a change in the potential barriers that the molecules have to surmount in order to reach the other potential minima. The change in the positions of the potential minima does not depend on temperature, and the polarization caused by it can be combined with that for atoms and electrons. The dipole polarization which is due to the passage of the hydrogen nuclei over the potential barriers is controlled by the barriers.

$$(\Delta U-\tfrac{1}{2}Ee\delta) \quad \text{and} \quad (\Delta U+\tfrac{1}{2}Ee\delta), \tag{5}$$

where E is the field, e the electronic charge and δ the separation of the two wells. When we establish a field E and change the polarization, there will be an internal field of the Lorentz type given by the equation

$$F=E+\beta P, \tag{6}$$

where β is $4\pi/3$ for an isotropic medium, but may differ from it for a crystal. The total polarization consists of a part P_E due to electrons and atoms, which includes the hydrogen potential well displacement polarization, and a part P_d due to hydrogen bond dipoles. The electronic and atomic polarization excited will be proportional to the local field F, so that

$$F=E+\beta[P_E+P_d]=E+\beta[\gamma F+P_d]$$

or

$$F=(E+\beta P_d)/(1-\beta\gamma) \tag{7}$$

where γ is the polarizability per unit volume due to all polarization except that of the hydrogen dipole. Hence α_+ and α_- then become, since δ lies at an angle of 25° with the applied field E_X,

$$\alpha_+=\Gamma\exp-\left[\Delta U-\frac{e\delta\cos(25°)}{2(1-\beta\gamma)}(E+\beta P_d)\right]/kT;$$

$$\alpha_-=\Gamma\exp-\left[\Delta U+\frac{e\delta\cos(25°)}{2(1-\beta\gamma)}(E+\beta P_d)\right]/kT.$$

Since $N_-=N-N_+$; $P_d=(2N_+-N)\mu$ and

$$dP_d/dt=\tfrac{1}{2}N\mu(\alpha_+-\alpha_-)-\tfrac{1}{2}P_d(\alpha_++\alpha_-) \tag{8}$$

TABLE I. Height of the potential barrier and distance of the H nucleus from the center of the bond as functions of the oxygen separation.

Length of bond in angstrom units	2.592	2.60	2.65	2.70	2.75	2.80
Height of potential barrier ΔU in calories per mole	0	15	314	975	1980	3310
Distance of H nucleus from center of bond in angstrom units ($\delta/2$)	0	0.058	0.157	0.217	0.265	0.308

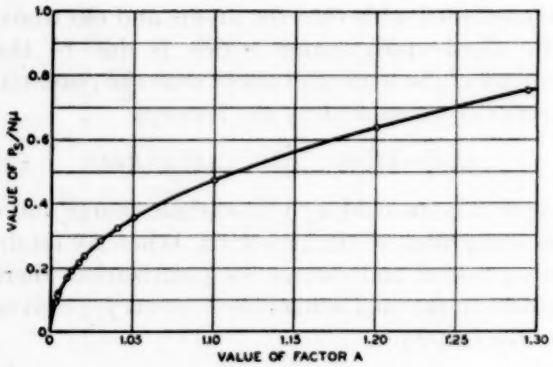


FIG. 3. Ratio of spontaneous polarization P_s to maximum dipole polarizations $N\mu$ as a function of A.

where μ is the dipole moment along the X axis, i.e., $\mu_A \cos(25°)$ where μ_A is the value of the dipole moment which lies in the direction of δ. Introducing the value of α_+ and α_- we have

$$\Gamma^{-1}e^{\Delta U/kT}\frac{dP_d}{dt}=N\mu\sinh\left[\frac{e\delta\cos(25°)}{2(1-\beta\gamma)kT}(E+\beta P_d)\right]$$

$$-P_d\cosh\left[\frac{e\delta\cos(25°)}{2(1-\beta\gamma)kT}(E+\beta P_d)\right]. \quad (9)$$

This is the fundamental equation determining all the dielectric properties of the crystal. To simplify writing, we introduce the abbreviation

$$A=\frac{e\delta N\mu\beta\cos(25°)}{2(1-\beta\gamma)kT}=\frac{e\delta N\mu_A\beta\cos^2(25°)}{2(1-\beta\gamma)kT}, \quad (10)$$

so that Eq. (9) can be written in the form[6]

[6] It has been shown by J. M. Richardson that even if we regard the atoms 1 and 10 as not quite equivalent, which will divide the bonds into two sets having dissymetrical potential barriers along the X axis, as shown by Fig. 2B, the form of Eq. (11) is still the same. If the two barriers have potentials U_1 and U_2 equal to $U_1=U+\Delta$ and $U_2=U-\Delta$, the polarization P_d consists of two parts Pd_1 and Pd_2 given for static conditions by the equations

$$Pd_1/N\mu=\tanh\left[\frac{A}{\beta N\mu}(E+\beta P_d)+\frac{\Delta}{kT}\right]$$

and

$$Pd_2/N\mu=\tanh\left[\frac{A}{\beta N\mu}(E+\beta P_d)-\frac{\Delta}{kT}\right].$$

On expanding the polarization in a power series, we have:

$$P_d/N\mu=(Pd_1+Pd_2)/N\mu$$
$$=[1-(\Delta/kT)^2][AP_d/N\mu]-\tfrac{1}{3}A^3(P_d/N\mu)^3.$$

Since Δ/kT is a small quantity, this results practically in defining a new value of A equal to

$$A'=A[1-(\Delta/kT)^2].$$

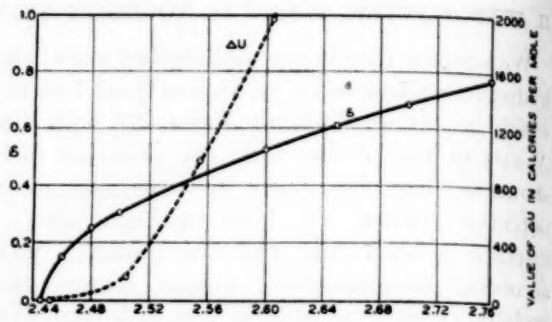


FIG. 4. Potential well separation δ and height of energy barrier ΔU as a function of the oxygen separation.

$$\Gamma^{-1}e^{\Delta U/kT}\frac{dP_d}{dt}=N\mu\sinh A\left[\frac{E}{\beta N\mu}+\frac{P_d}{N\mu}\right]$$

$$-P_d\cosh A\left[\frac{E}{\beta N\mu}+\frac{P_d}{N\mu}\right]. \quad (11)$$

For static conditions where $dP/dt=0$, this becomes

$$\tanh A\left[\frac{E}{\beta N\mu}+\frac{P_d}{N\mu}\right]=P_d/N\mu. \quad (12)$$

Setting the field equal to zero, this equation becomes

$$P_d/N\mu=\tanh(AP_d/N\mu). \quad (13)$$

If the factor A is greater than unity, this equation will have positive and negative solutions other than zero, representing spontaneous polarization along the $+X$ or $-X$ direction. For values of A slightly greater than unity, we can replace $\tanh(AP_d/N\mu)$ by the first two terms of the expression or

$$(P_d/N\mu)=A(P_d/N\mu)-\tfrac{1}{3}A^3(P_d/N\mu)^3.$$

Solving for $P_d/N\mu$, we have

$$(P_d/N\mu)=(3(A-1)/A^3)^{\frac{1}{2}}. \quad (14)$$

A plot of this equation is given in Fig. 3.

For rochelle salt with cell dimensions 14.3A $\times$11.93A$\times$6.17A, the unit cell has a volume of 1.05×10^{-21} cm^3. Since there are four dipoles per unit cell, the number N is 3.81×10^{21} dipoles per cm^3. If the hydrogen nucleus is midway between the two oxygens the bond is neutral and there is no dipole moment. Hence the value of the dipole moment should be given by the expression

$$\mu_A=\tfrac{1}{2}e\delta \quad (15)$$

and the expression for A becomes

$$A = e^2\delta^2\beta N \cos^2(25°)/4(1-\beta\gamma)kT. \quad (16)$$

In order to calculate A we have to know how the separation δ depends on the separation of the oxygens 1 and 10. Some idea of the change in δ can be had from a calculation of Huggins[7] on the shape of the potential curve for a free hydrogen bond. He finds that the shape of the curve depends on the oxygen separation. Table I shows the height of the potential barrier and the distance of the H nucleus from the center of the bond ($\delta/2$) as a function of the separation of the oxygens. Unfortunately, with the constants chosen, these data do not agree with the accepted values for water which, according to Pauling,[8] are a separation of 2.76A results in a value of $\delta=0.78$A. If we adjust the constants of Huggin's equation to give this value, the values of δ and ΔU are shown plotted on Fig. 4.

As an indication of the consistency of this theory, a calculation of the value of A is attempted at 297°K, the upper Curie point where A should be unity. From x-ray data, the value of the oxygen separation is 2.59 ± 0.05A. All the experimental data are found to be satisfied best if we take the separation as 2.57A, a value within the experimental error. With this value $\delta=0.4655$A. All the quantities of Eq. (16) are known except β and γ the polarizability. An estimate of these can be made as follows. The values of β and γ are related to the dielectric constant for electrons and atoms by the formula

$$(\epsilon_0-1)/[3+3\beta(\epsilon_0-1)/4\pi]=4\pi\gamma/3, \quad (17)$$

where ϵ_0 is the dielectric constant for electrons and atoms, i.e., due to all sources except the hydrogen bond dipoles. Since the index of refraction for rochelle salt along the X axis is 1.4954, the electronic dielectric constant is the square of this, or 2.24. At −160°C, the dielectric constant is 7.0. As discussed later, this dielectric constant, which is independent of temperature for temperatures lower than −160°C, is due to the electronic and atomic polarization plus a polarization due to the displacement of the hydrogen nuclei due to the change in the equilibrium position of a single potential well. For the double potential well of Fig. 2A, the polarization due to the displacement of the equilibrium positions will be half as large so that a value $\epsilon_0=6.0$ is assumed, which seems to be a reasonable value. Substituting in Eq. (17)

$$\gamma=1/(\tfrac{1}{3}4\pi+\beta). \quad (18)$$

Substituting all the known values in the expression for A, we have

$$A=\frac{(4.77\times10^{-10})^2(0.465\times10^{-8})^2\times3.81\times10^{21}\times0.819\beta}{4\times1.37\times10^{-16}\times297\left[\dfrac{4\pi/5}{4\pi/5+\beta}\right]}=1. \quad (19)$$

The only unknown is β and solving for this, $\beta=4.07$ which is very close to the theoretical value for an isotropic substance $4\pi/3=4.19$.

To calculate A over a temperature range, one would have to know how the separation of the oxygen atoms varied with temperature on which no data are available. However, since the hydrogen bond is the weakest bond in the structure, one would expect that a good share of the temperature expansion of the crystal would come across this bond. Measurements[9] of the temperature expansion along the X axis show that

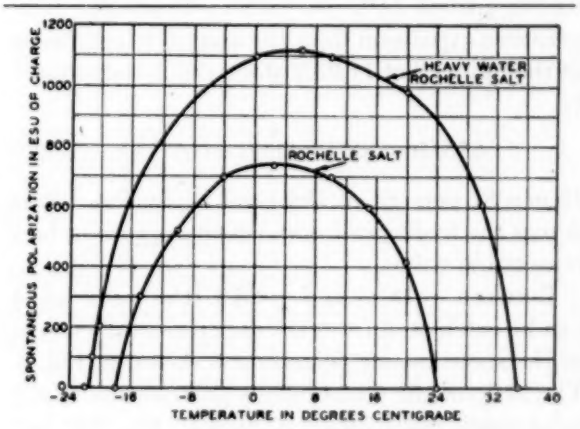


FIG. 5. Spontaneous polarization of rochelle salt and heavy water rochelle salt plotted as a function of temperature.

[7] M. L. Huggins, J. Phys. Chem. **40**, 723 (1936).
[8] L. Pauling, *Nature of the Chemical Bond* (Cornell University Press, Ithaca, New York, 1945), Chapter IX, plate 361.
[9] W. P. Mason, Bell Sys. Tech. J. **26**, No. 1 (1947).

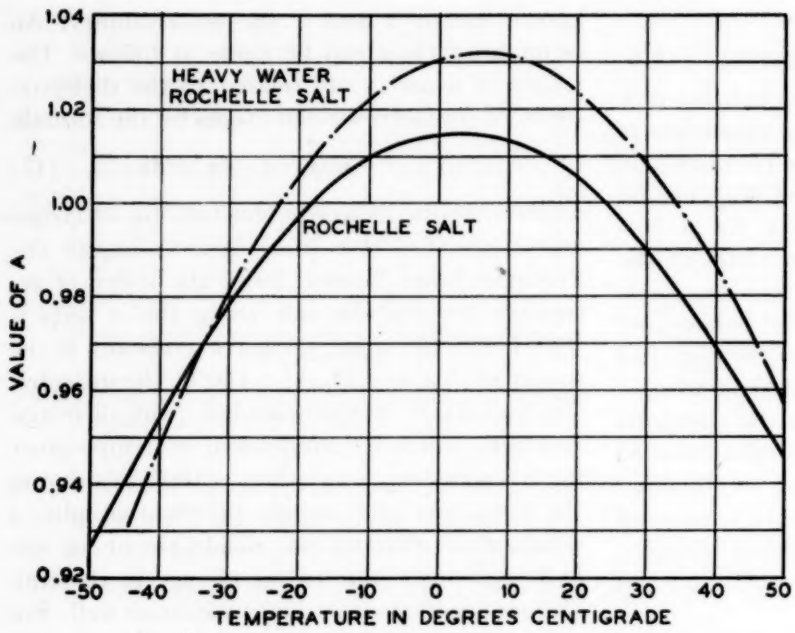


FIG. 6. Calculated value of A for rochelle salt and heavy water rochelle salt as a function of temperature.

from -50°C to $+40$°C, the temperature expansion averages about 65×10^{-6} per degree C. Since the value of $A=1$ at -18°C, the other Curie temperature, an estimate can be made of the proportion of the total expansion that comes across the hydrogen bond. The values of N, β, and γ will all vary slightly with temperature but their variations should be small compared to δ and hence it will be assumed that the complete variation occurs in δ. To obtain $A=1$ at -18°C requires a value of $\delta=0.431$A. From Fig. 4 this corresponds to an oxygen separation of 2.5535A. This represents a decrease in separation of 0.0165A from 24°C to -18°C. From the temperature expansion data the complete expansion of the unit cell of 11.93A along the X axis should be

$$65\times10^{-6}\times42\times11.93=0.0325\text{A}. \qquad (20)$$

Hence 51 percent of the total expansion occurs across the hydrogen bond, which appears to be a reasonable value.

Between -18°C and $+24$°C the value of A can be calculated from the experimental[10] values of the spontaneous polarization shown in Fig. 5. The ratio of the spontaneous polarization to the total dipole polarization is given by Eq. (14) and is plotted in Fig. 3. The total dipole polarization along the X axis at 24°C (297°K) is

$$N\mu=\tfrac{1}{2}Ne\delta\cos(25°)=\tfrac{1}{2}3.81\times10^{21}\times4.77\times10^{-10}\times0.4655\times10^{-8}\times0.904=3850\text{ esu}.$$

At -18°C (255°K), since δ has decreased to 0.431A, $N\mu=3550$ esu. Assuming that $N\mu$ is a straight line function of temperature between these points, the value of A can be calculated between -18°C and $+24$°C and is shown plotted on Fig. 6. It is very closely a parabola, and hence we can extend the values on either side of the Curie region as shown in Fig. 6. From the values of A, the values of δ can be calculated and from Fig. 4 the separation of the oxygens can be calculated. These are shown plotted on Fig. 7. If we extrapolate the curves to low temperatures it appears that a separation of 2.446A (the separation for which $\delta=0$ and the two potential wells disappear) will occur at about -160°C. At this temperature, the temperature variable part of the dielectric constant should disappear. Some confirmation[11] of this calculation is found experimentally for at that temperature the dielectric constant is 7.0 and does not change as the temperature is lowered. Figure 7 indicates that the separations of the oxygens increases more slowly with temperature for temperatures

[10] H. Mueller, Ann. Acad. Sci. N. Y. 40, 338 (1940).

[11] See W. G. Cady, *Piezoelectricity* (McGraw-Hill Book Company, Inc., New York, 1946), p. 559, Fig. 121.

above -30°C. This may be accounted for by the small mass of the hydrogen nucleus which acquires its full kinetic energy at a lower temperature than do the heavier molecules of the crystal. At the higher temperatures these other atoms acquire their full motion and hence the expansion along the hydrogen bond becomes less in proportion to the total expansion.

Equation (12) and the constants evaluated from the above consideration give methods for determining the properties of the crystal at large field strengths. This equation holds primarily for a single domain. Since $(A/\beta N\mu)$ has a value of 6.7×10^{-5} at 0°C, the hyperbolic tangent of Eq. (12) can be expanded into the form

$$\tanh A\left[\frac{E}{\beta N\mu}+\frac{P_d}{N\mu}\right]$$

$$\frac{\tanh(AE/\beta N\mu)+\tanh(AP_d/N\mu)}{1+\tanh(AE/\beta N\mu)\tanh(AP_d/N\mu)}$$

$$\doteq 6.7\times10^{-5}E+\tanh AP_d/N\mu. \quad (21)$$

Since $AP_d/N\mu$ is in the order of 0.20 at the largest, we can replace the hyperbolic tangent by the first two terms of its power series, and Eq. (21) can be written

$$(AP_d/N\mu)-\tfrac{1}{3}A^3(P_d/N\mu)^3 = -6.7\times10^{-5}E+P_d/N\mu. \quad (22)$$

If P_s is the solution for the spontaneous polarization when $E=0$, we can write

$$P_d=P_s+\Delta P.$$

Solving for ΔP, we find

$$\Delta P=(AE/\beta)/[A^3(P_s/N\mu)^2-(A-1)]. \quad (23)$$

Since $[4\pi(\Delta P+P_E)+E]$ is the electric displacement for fields above the saturation effect, the dielectric constant for values of the field so high that domain saturation has set in will be

$$\epsilon_s=D/E=\epsilon_0 +4\pi A/\beta[A^3(P_s/N\mu)^2-(A-1)]. \quad (24)$$

At 0°C, $P_s=740$ e.s.u., $N\mu=3700$ e.s.u., $A=1.014$ and $\beta=4.07$. Hence the saturation dielectric constant for the clamped crystal should have a value of 114. This agrees reasonably well with the measured value[12] of 200 at 20°C, where the saturation dielectric constant should be somewhat higher due to the smaller denominator of Eq. (24).

The theoretical value of the coercive field required to reverse a single domain can also be calculated from Eq. (22). If we plot the curve on a graph, the shape is shown in an exaggerated form in Fig. 8. To determine the coercive voltage required to change the polarization to the opposite sign, we have to raise E (negative) until

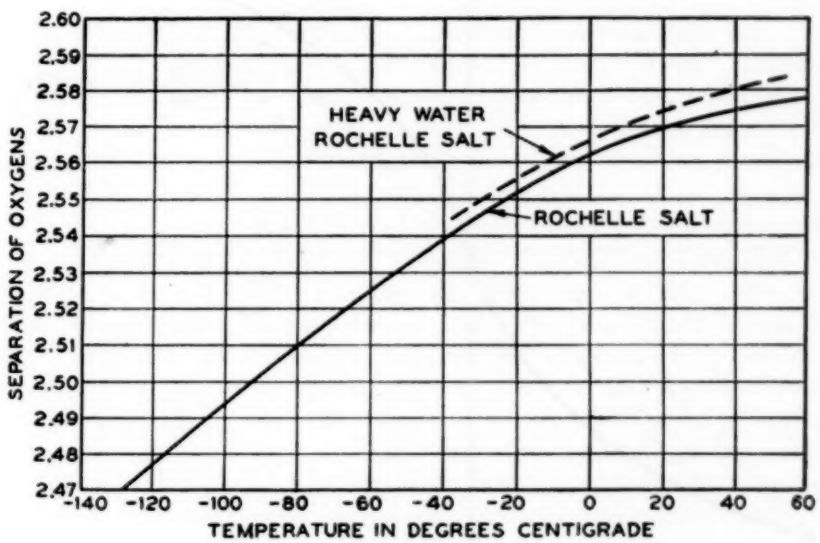


FIG. 7. Calculated separation of oxygens in rochelle salt and heavy water rochelle salt as a function of temperature.

[12] Reference 1, p. 569.

-6.7×10^{-5} E equals the greatest deviation of the hyperbolic tangent from the straight line going through zero and the solutions $P_d/N\mu = 0.20$. By calculation, this occurs when $P_d/N\mu = 0.12$ and it requires a coercive field of 16 esu or 4800 volts cm. As with the phenomenological theory, this is about 16 times the experimental value. An explanation of this discrepancy has been given by Cady.[13] The effect is caused by the other domains contributing a field determined by the applied electric field, which cuts down the value of the applied field necessary to reverse the domain.

III. CLAMPED DIELECTRIC CONSTANT AT LOW FIELD STRENGTHS

A solution of Eq. (11) for small applied voltages also accounts quantitatively for the measured values of the clamped dielectric constant at low field strengths over a frequency and temperature range. For simple harmonic motion we let

$$E=E_0e^{j\omega t};\quad P_d=P_s+P_0e^{j\omega t}$$

where P_s is the spontaneous dipole polarization. Also since the time-variable field and polarization are small compared to the spontaneous effect, we can write

$$\sinh A\left[\frac{E_0e^{j\omega t}}{\beta N\mu}+\frac{(P_s+P_0e^{j\omega t})}{N\mu}\right]$$
$$=\sinh A\left[\frac{E_0}{\beta N\mu}+\frac{P_0}{N\mu}\right]e^{j\omega t}\cosh\frac{AP_s}{N\mu}$$
$$+\cosh A\left[\frac{E_0}{\beta N\mu}+\frac{P_0}{N\mu}\right]e^{j\omega t}\sinh\frac{AP_s}{N\mu}$$

and since the time variable parts are very small

$$\sinh A\left[\frac{E_0+\beta P_0}{\beta N\mu}\right]e^{j\omega t}\doteq\frac{A[E_0+\beta P_0]e^{j\omega t}}{\beta N\mu};$$
$$\cosh A\left[\frac{E_0+\beta P_0}{\beta N\mu}\right]e^{j\omega t}\doteq 1$$

Then Eq. (11) takes the form

$$\Gamma^{-1}e^{\Delta U/kT}j\omega P_0e^{j\omega t}$$
$$=N\mu\left[\sinh\frac{AP_s}{Nu}+\frac{A(E_0+\beta P_0)}{\beta N\mu}e^{j\omega t}\cosh\frac{AP_s}{N\mu}\right]$$
$$-(P_s+P_0e^{j\omega t})\left[\cosh\frac{AP_s}{N\mu}+\frac{A(E_0+\beta P_0)}{\beta N\mu}\right.$$
$$\left.\times e^{j\omega t}\sinh\frac{AP_s}{N\mu}\right]. \quad (26)$$

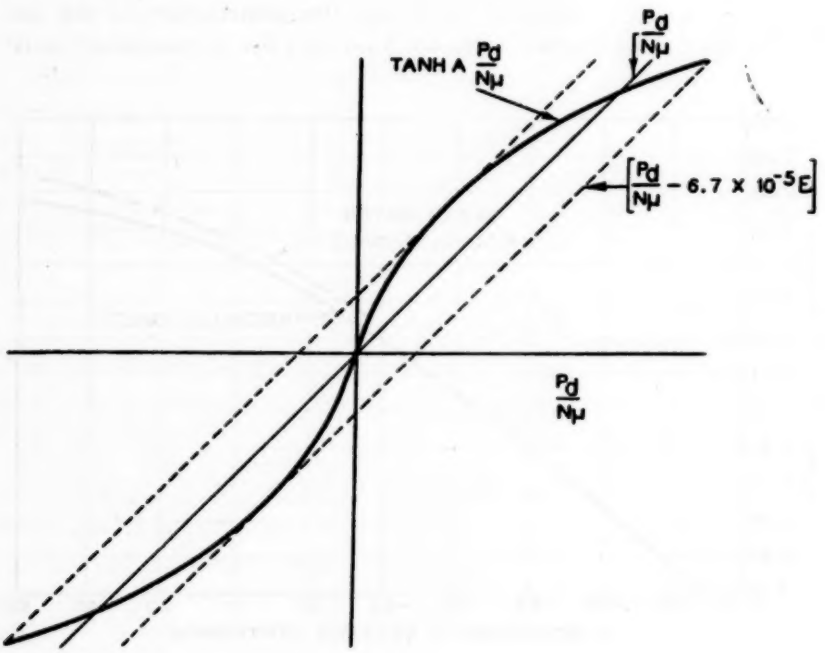


FIG. 8. Value of Eq. (22) plotted as a function of $P_d/N\mu$.

[13] Reference 1, p. 654.

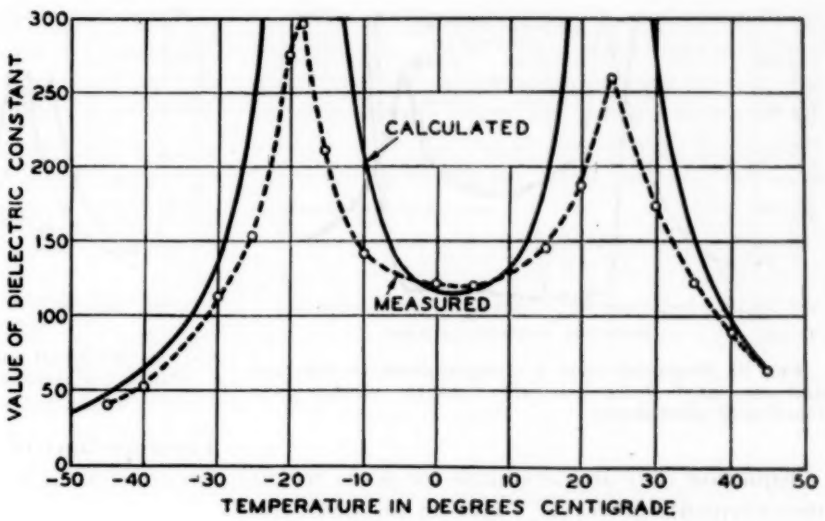


FIG. 9. Measured and calculated values of the clamped dielectric constant for low field strengths plotted as a function of temperature.

Collecting the constant terms we have again the equation for spontaneous polarization

$$\tanh(AP_s/N\mu)=P_s/N\mu, \tag{13}$$

while the time variable polarization can be written in the form

$$P_0=\frac{(A/\beta)[1-(P_s/N\mu)^2]E_0}{1-A[1-(P_s/N\mu)^2]+j\omega e^{\Delta U/kT}/\Gamma\cosh(AP_s/N\mu)}. \tag{27}$$

The electric displacement D is equal to

$$D=4\pi P_0+4\pi P_E+E=\epsilon E \tag{28}$$

where P_E is the electronic and atomic polarization and ϵ the dielectric constant. Hence the dielectric constant of the clamped crystal for small applied fields takes the form

$$\epsilon=\epsilon_0+\frac{(4\pi A/\beta)[1-(P_s/N\mu)^2]}{1-A[1-(P_s/N\mu)^2]+j\omega e^{\Delta U/kT}/\Gamma\cosh(AP_s/N\mu)}. \tag{29}$$

where ϵ_0 is the dielectric constant for electrons and atoms alone which has been assumed to be $\epsilon_0=6.0$.

The measured value of the spontaneous polarization is given by Fig. 5. From this and the calculated value of A given by Fig. 6, the calculated value of the clamped dielectric constant of rochelle salt is found as shown by Fig. 9. This quantity was measured at frequencies of about 160 kilocycles by measuring[14] the longitudinally clamped dielectric constant (i.e., the dielectric constant when longitudinal motion is suppressed), and correcting for the remaining motion. This was measured from -10°C to $+50$°C and is shown plotted by the solid line of Fig. 10. In order to eliminate any corrections, and to extend the temperature range, the dielectric constant of the clamped crystal has been measured by measuring the dielectric constant of an X-cut rochelle salt crystal at 20 megacycles. The dimensions of the crystal were 1.75 cm by 1.75 cm in cross section and 0.759 cm thick (thickness along X axis). This is so high a frequency that the added capacity due to any piezoelectric resonances is negligible and the measured capacity is directly the clamped dielectric constant. The measured values of the dielectric constant and the Q of the crystal are shown plotted by the dashed line. The measured values agree quite well with those obtained previously. The points of highest capacity come at -18°C and $+24$°C which, as mentioned in the introduction, show conclusively that the anomaly must reside in the clamped dielectric constant. The measured values are somewhat lower than the calculated values obtained from Eq. (29) using the values of A from Fig. 6, the spontaneous polarization from Fig. 5, the value of $N\mu=3850$ e.s.u. at 24°, 3550 at -18°, and a proportional value for any other temperature, a value of $\beta=4.07$ and a value of $\epsilon_0=6$.

[14] W. P. Mason, Phys. Rev. **58**, 747 (1940).

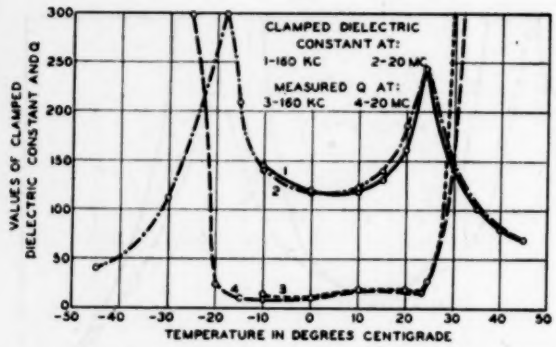

CLAMPED DIELECTRIC CONSTANT AT:
1-160 KC 2-20 MC
MEASURED Q AT:
3-160 KC 4-20 MC
VALUES OF CLAMPED DIELECTRIC CONSTANT AND Q
TEMPERATURE IN DEGREES CENTIGRADE

FIG. 10. Measured value of clamped dielectric constant and associated Q for low field strengths plotted as a function of temperature.

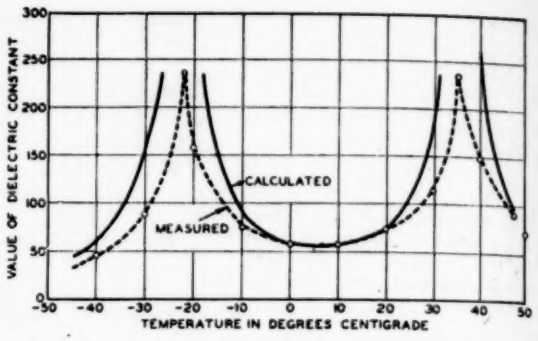


FIG. 11. Measured and calculated values of the clamped dielectric constant of heavy water rochelle salt.

Equation (29) also accounts for some recent measurements of W. A. Yager on the dielectric constant of rochelle salt at a frequency of 2.5 $\times 10^{10}$ cycles. Yager finds that from $-40°$C to $+26°$C the dielectric constant and the Q of rochelle salt are practically independent of temperature and have the values

$$\epsilon = 8; \quad Q = 0.25. \tag{30}$$

These measurements were made by a wave-guide technique. They indicate that the frequency is so high that the hydrogen bond dipoles cannot follow the field, and contribute little to the dielectric constant.

The impedance to be expected theoretically can be calculated from Eq. (29). This results in

$$\frac{1}{Z} = \frac{1}{Z_d} + \frac{1}{Z_0} = j\omega \frac{\epsilon \times 1.11 \times 10^{-12}}{4\pi} = j\omega \times 1.11 \times 10^{-12} \times \left[\frac{(A/\beta)[1-(P_s/N\mu)^2]}{1-A[1-(P_s/N\mu)^2] + j\omega e^{\Delta U/kT}/\Gamma \cosh(AP_s/N\mu)} + \frac{\epsilon_0}{4\pi} \right].$$

This equation is the impedance of two parallel arms, one arm Z_d representing the impedance of the dipole term and the other, the impedance Z_0, is the impedance of a condenser with a dielectric constant ϵ_0 due to other types of polarization. The impedance of the dipole arm per cubic cm is

$$Z_d = 9 \times 10^{11} \left[\frac{\beta}{A(1-(P_s/N\mu)^2)} \right] \left\{ \left[\frac{\Gamma^{-1} e^{\Delta U/kT}}{\cosh(AP_s/N\mu)} \right] - \frac{j}{\omega} [1 - A(1-(P_s/N\mu)^2)] \right\}. \tag{31}$$

The impedance of the other arm with the dielectric constant ϵ_0 is

$$Z_0 = -\frac{j}{\omega C_0} = -\frac{j9 \times 10^{11} \times 4\pi}{\omega \epsilon_0}. \tag{32}$$

In evaluating this equation we now assume the value of Γ given by Eyring's reaction rate theory, namely $\Gamma = (kT/h)$. At a frequency of 2.5×10^{10} cycles and at a temperature of 0°C, the impedance of the dipole arm, assuming $\beta = 4.07$;

$$A = 1.014, \quad (P_s/N\mu) = 0.20, \quad h = 6.55 \times 10^{-27}, \quad k = 1.37 \times 10^{-16}, \quad T = 273, \quad \Delta U = 1100$$

calories per mole, is

$$Z_0 = 3.2 - j0.63 \tag{33}$$

and hence the resistance component is larger than the reactance component. The indicated relaxation frequency is about 5×10^9 cycles at 0°C. Furthermore we see that the value of the resistance will not change much with temperature since $(\Delta U/kT)$ decreases as we decrease the temperature and this offsets the increase due to $1/kT$. Hence the measured capacitance shunted by the resistance will be relatively independent of the temperature over the region $-40°$C to $+20°$C. The measured results of Yager given by Eq. (30) indicate that the impedance per cm³ at 2.5×10^{10} cycles, should be a reactance $-j9.0$ ohms in parallel with a resistance $\frac{1}{4}$ this value, or 2.25 ohms. Neglecting the reactance $-j0.63$ in Eq. (33) the resistance value is within a factor of 2 of the measured value and indicates that the best separation of oxygens is about 2.57A. The indicated dielectric constant of 8 is somewhat larger than the value 6 assumed, but

since the value 8 is a preliminary result it agrees reasonably well with the value 6.0 assumed. Hence the present theory accounts approximately for the clamped dielectric constant and the spontaneous polarization for all temperatures and frequency ranges. It involves only molecular constants that can in principle be determined by x-ray measurements, rather than the phenomenological constants.

The present theory does not account for the finite value of the clamped dielectric constant at the Curie points nor the hysteresis resistance at small amplitudes both inside and outside of the Curie temperatures, which is evidenced by the nearly constant values of Q, *versus* frequency, shown by the curves of Fig. 10. It is thought that these effects may be due to the fact that the hydrogen bond 1 to 10 may be part of a chain of three hydrogen bonds. When a displacement occurs in the 1 to 10 bond it may generate coupled displacements for the connecting bonds which react back on the 1 to 10 bond. This would cause a hysteresis loop since it would require a reverse force to undo the effect of the connecting link displacement and the loss of energy to the connecting bonds may also be the cause of the finite dielectric constant at the Curie temperatures.

IV. APPLICATION OF THEORY TO HEAVY WATER ROCHELLE SALT

When the hydrogens in the water molecules and the hydroxyl molecules are replaced by heavy hydrogens, a very considerable change occurs in the properties of the rochelle salt crystal. As shown[15] by Fig. 5, the lower Curie point is lowered to -22°C, the upper one raised to 35°C and a considerably larger spontaneous polarization occurs. The properties of heavy water rochelle salt provide a good test of the present theory. They are accounted for if we assume that the separation of the oxygens is increased in the heavy water rochelle salt over ordinary rochelle salt.

From the ratio of the spontaneous polarization 1120 e.s.u. to the value of $N\mu$ which is still about 3700 e.s.u., and the curve of Fig. 3, one can see that the replacement of hydrogen by heavy hydrogen has resulted in the factor

$$A = e^2\delta^2 N\beta \cos^2 25°/4kT(1-\beta\gamma) = 1.033,$$

at 6°C for which the highest value of A occurs. To account for this value and $A=1$ at -22°C and $+35$°C, the oxygen separation *versus* temperature occurs as shown by the dashed line of Fig. 7. In general this separation parallels that for the ordinary rochelle salt, but at the higher temperature shows less bending off than does that for ordinary rochelle salt. This is probably accounted for by the double weight of the heavy hydrogen nucleus. Using these values of separations, the values of A for heavy water rochelle salt are shown plotted by the dot-dash line of Fig. 6.

From the value of A given by Fig. 7 and the measured value of the spontaneous polarization shown by Fig. 5, the calculated clamped dielectric constant is shown plotted by Fig. 11. The measured values can be obtained from previous measurements[16] of the free dielectric constant and the piezo-electric constant, and are shown plotted by the dotted line of Fig. 11. This agrees well with the calculated value.

[15] J. Hablutzel, Helv. Phys. Acta. **12**, 489 (1939).

[16] A. N. Holden and W. P. Mason, Phys. Rev. **57**, 54 (1940). J. Hablutzel, Helv. Phys. Acta. **12**, 489 (1939).

Resonance Scattering of Neutrons by Cobalt*

S. P. HARRIS, A. S. LANGSDORF, JR., AND F. G. P. SEIDL
Argonne National Laboratory, Chicago, Illinois

(Received September 17, 1947)

THE discovery that the Mn^{55} neutron resonance[1] about 300 ev is predominantly scattering[2-4] has induced us to search for other similar resonances. We have studied Co^{59} and conclude that the Co neutron resonance[5] near 115 ev is also mainly scattering.

If the neutron width is much greater than the absorption width, then according to the Breit-Wigner equation the peak scattering cross section for Co^{59} (spin 7/2, resonance 115 ev) should be 10,000 or 13,000 b ("b" is equivalent to 10^{-24} cm^2), depending upon the spin of the compound nucleus. Experimentally, we find a total peak cross section of at least 6000 b.

The transmission of a layer of $CoSO_4$ solution in D_2O (0.0233 g/cm^2 Co) was measured for neutrons detected by the 10.7-min. activity which they produced in a thin electrolytic Co foil (0.01 g/cm^2). The detector foil was located 200 cm behind the $CoSO_4$ solution and aligned in a collimated neutron beam from the heavy-water pile. The beam was passed through B^{10} and thin Cd to remove slow neutrons. Accordingly, the average total cross section was found to be 3100±500 b for neutrons filtered through 0.15 g/cm^2 B^{10} and 2700±500 b for the case of 0.31 g/cm^2 B^{10}. A large counting error resulted here because of the low counting rate caused by the softness of the 10.7-min. Co^{60} radiations.

In a second experiment, neutrons scattered by a Co foil were detected by an annular BF_3 counter surrounding the foil. The counting region lay between an outer Al tube 4 in. O.D. and a concentric inner Al tube $1\frac{7}{8}$ in. I.D. These were welded to suitable end rings so that the smaller tube formed an open path through the counter. The collecting electrode consisted of a cylindrical grid of 2-mil wires mounted coaxially within the annulus. The assembly was filled with pure BF_3 to 15-cm Hg and operated at about 1900 v as a proportional counter. When in use, a collimated pencil of neutrons was directed down the $1\frac{7}{8}$-in. tube through the system, and the scattering sample was placed in the beam near the center of this tube. The efficiency for detecting scattered neutrons was increased by surrounding the counter with a paraffin reflector, except in the path of the beam. To reduce background, the reflector was covered by Cd and by a shield of B_2O_3 and paraffin. Elimination of air scattering, by closing off the ends of the $1\frac{7}{8}$-in. tube with Al foils and evacuating, further decreased the background by a factor of ten.

In the second experiment an electrolytic Co foil was used as a scatterer. Near the Co resonance the efficiency of counter and reflector change relatively slowly with neutron energy; therefore, in this region the over-all sensitivity of Co scatterer, reflector, and counter should vary as the scattering cross section of Co. By means of this detector system the transmission of another Co foil (0.00677 g/cm^2) was measured. The thinner foil was inserted in the beam 15 in. in front of the counter, and an auxiliary collimator was placed between, so that practically

TABLE I. Summary of measurements.

Filters in neutron beam				Transmission T	
Cd g/cm^2	B^{10} g/cm^2	E_e* ev	E_m** ev	of 0.0067 g/cm^2 of Co	σ_{eff}*** b
0.00	0.000	—	—	0.919±0.003	1220± 40
0.34	0.000	—	—	0.850±0.006	2360± 90
0.34	0.157	37	8.4	0.852±0.005	2320± 85
0.34	0.313	163	33.6	0.869±0.009	2030±150
0.34	1.267	5600	550	0.976±0.026	360±360

* A neutron of energy E_e has the probability $1/e$ of penetrating the corresponding B^{10} thickness.
** Assuming a $1/E$ epi-Cd incident neutron spectrum, E_m is the most probable energy of a neutron traversing the corresponding B^{10} thickness.
*** $\sigma_{eff} = \ln(1/T)/\tau$, T being the transmission of a layer of thickness τ atoms/cm^2.

* This research has been carried out under contract No. W-31-109-eng-38 for the Atomic Energy Commission.

[1] L. J. Rainwater, W. W. Havens, Jr., C. S. Wu, and J. R. Dunning, Phys. Rev. 71, 65 (1947).
[2] M. Goldhaber and A. A. Yalow, Phys. Rev. 69, 47A (1946).
[3] N. H. Barbre and M. Goldhaber, Phys. Rev. 71, 141A (1947).
[4] F. G. P. Seidl, S. P. Harris, and A. S. Langsdorf, Jr., Phys. Rev. 72, 168A (1947).
[5] C. S. Wu, L. J. Rainwater, and W. W. Havens, Jr., Phys. Rev. 71, 174 (1947).

..one of the neutrons scattered by the first foil would be detected. The results appear in Table I, where the peak cross section should be at least twice σ_{eff}, which is an average over the resonance.

By the use of these data, the effect in the neighborhood of the Co resonance may be accentuated by computing the transmission for neutrons passed by 0.157 g/cm² B^{10}, but not by 0.313 g/cm² B^{10}; most such neutrons have energies between 2 and 200 ev. In this way $\sigma_{eff}=2920 \pm 150$ b, and an approximate integration shows that this value is consistent with a total *peak* Co cross section of 8000 to 9000 b.

Also, when the neutron transmission of thin Mn was investigated by means of scattering by a Co foil, the σ_{eff} for Mn was found to be ~ 10 b, and similarly for the transmission of Co as measured by Mn scattering. Thus resonance characteristics were exhibited by Mn and Co.'

The annular counter was used in a third experiment to compare the scattering by Co, Mn, and graphite (1) for the total pile flux, including thermal neutrons, and (2) for neutrons filtered by Cd. If we assume the graphite-scattering cross section remains constant at 4.8 b, a large excess of epi-Cd scattering is apparent for both Co and Mn. Associating the excess scattering with a single level, we have roughly estimated a lower limit for $\sigma_0\Gamma$, where σ_0 denotes the peak *scattering* cross section and Γ the total width. In this calculation the sensitivity of counter plus paraffin reflector was assumed to vary with neutron energy E in either of two ways: (a) constant from $E=0$ to 10^4 or 10^6 ev, and thereabove zero; (b) $1/\sqrt{E}$. For Co $\sigma_0\Gamma \cong 20{,}000$ b-ev or 58,000 b-ev, respectively, for case (a) or (b); while for Mn $\sigma_0\Gamma \cong 30{,}000$ b-ev or 140,000 b-ev. These values are evidence of a very large resonance scattering.

Letters to the Editor

PUBLICATION of brief reports of important discoveries in physics may be secured by addressing them to this department. The closing date for this department is five weeks prior to the date of issue. No proof will be sent to the authors. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents. Communications should not exceed 600 words in length.

Temperature Effect of Cosmic Radiation at 1000-M Water Equivalent Depth

M. FORRÓ
Institute for Experimental Physics, University of Budapest, Budapest, Hungary
September 15, 1947

IN the course of experiments performed together with J. Barnóthy in a coal mine at 1000-m water-equivalent depth, I had occasion to note the intensities of the most penetrating component of cosmic radiation and the outer-air temperature for a period of three years. Temperature correlation was not the main purpose of the experiments; accordingly, a number of different arrangements were used.

The coal mine is situated at 35-km NW from Budapest. The apparatus was installed at 470-m depth below the surface in a side gallery of the mine, the latter having a diameter of 2 m, with a brick casing of 1-m thickness. The temperature at this location was constant within ±1°C during the whole year.

The intensities were measured with a threefold coincidence-counter train of 8×100 cm² sensitive area. Each tray consisted of two parallel connected G-M counters. The coincidences were selected by means of a Barnóthy circuit, having a resolving time of 1.4×10^{-6} sec. (This circuit is essentially identical with that generally known as the Getting circuit; its description was published in Naturwissenschaften 21, 835 (1933).) The distance between extreme trays was 101.3 cm, and provision was made for interposing 80 cm of lead between the counter trays. The adopted procedure was to put the same amount of lead simultaneously between top and center counter tray and center and bottom counter tray. With every thickness of lead the intensities were recorded without interruption in average for two weeks, and the whole absorption curve was surveyed several times. The temperature correlation was reckoned for every thickness of lead, i.e., for every arrangement separately.

A total of 2066 coincidences was recorded during 10,226 hours. Column 1 of Table I indicates the total thickness of the interposed lead; column 2, the date; column 3, the total time of the measurement; column 4, the total number of counts; column 5, the mean outer-air temperature for the same period; column 6, the correlation coefficient derived from the different runs with the same arrangement; column 7, the temperature effect and its probable error computed with its help. In order to compute the temperature effect, the different lengths of the periods were considered.[1] Some of the individual values have relatively large errors, but the mean of the nine individual results yields a temperature effect: $+0.74\pm0.13$ percent per degree C, which value exceeds about six times its probable error and, therefore, the existence of a positive temperature effect is ascertained.

The temperature values indicated in the table were kindly supplied by the Hungarian Meteorologic Office; they refer to the records of the meteorologic station of Esztergom, situated about 8-km north from the mine. Unfortunately, no upper-air temperature data were available.

Performing a similar computation with barometric pressure, we obtain for the barometer effect:

$$BE=+0.42\pm0.90 \text{ percent per mm Hg.}$$

This means that no correlation exists between cosmic-ray intensity at great depth and barometric pressure. The absence of a barometer effect rendered the application of multiple correlation superfluous, and the observed temperature effect can in no case be attributed to an overlooked influence of barometric pressure. The analysis revealed that factors other than temperature do not influence the intensity.

The temperature effect obtained has the sign which we would expect if in accordance with the view expressed previously;[2] the most penetrating component of cosmic radiation consists of neutrini created by the disintegration of the mesons in the atmosphere. In this case the intensity of the neutrini, i.e., the number of neutrini with energies greater than E

$$I_\nu=\text{const.}\int_{2E}^{\infty}\frac{1-\exp(-L/l)}{E^{\gamma+1}},$$

where

$$l=\tau_0 E/\mu c$$

increases with increasing temperature, since if the place where the mesons are formed is shifted upwards, a larger number of mesons disintegrate before reaching sea level, and, consequently, a greater number of neutrini with great energies are created. I should like to emphasize that the conception that at great depth only meson-decay products are observable is supported by other observations. J. Barnóthy[3] has shown that according to this assumption the exponent γ, of the energy in the absorption law of the radiation, increases from 1.8 to 2.8 for depth below 300-m water equivalent, and, further, that at great depth—contrary to the results found by Auger at small depth—the intensity is relatively greater in inclined directions than would correspond to the thickness of the absorbing layer, in agreement with experimental findings. According to this interpretation we should have to expect for the radiation below 300-m water equivalent a positive temperature effect.

Let us assume that mesons are formed at a height of 80-mm Hg. According to aerologic observations 80-mm Hg corresponds in average to $L=15{,}935$ m, and it has a temperature variation with respect to the temperature at

TABLE I. Summary of results.

Lead in cm	Date	Total time in hours	Total number of counts	Mean temperature	Correl. coefficient	Temperature effect
0	1941 VI	662.7	165	20.7°	−0.702	−0.243±0.084
	1942 XII	477.1	125	0.1		
	1943 III	192.9	46	10.3		
10	1941 VII	466.5	94	20.2	+0.914	+1.015±0.077
	1941 IX	520.5	106	15.4		
	1943 II	343.8	59	4.6		
15	1941 IX	321.2	75	13.4	+0.939	+1.152±0.070
	1941 XII	383.4	74	−2.1		
	1943 II	356.4	78	2.9		
20	1941 X	230.0	55	5.6	+0.279	+0.861±1.35
	1941 XI	249.8	45	1.1		
	1943 I	478.0	106	−2.2		
25	1941 III	284.0	52	0.7	+0.950	+0.793±0.039
	1941 X	305.2	57	7.0		
	1943 I	359.2	61	−5.0		
35	1941 I	416.5	76	−2.7	+0.530	+0.400±0.26
	1941 V	366.4	68	18.3		
	1943 VI	398.0	85	16.5		
40	1940 XII	312.7	57	−8.0	+0.107	+0.082±0.36
	1941 X	344.7	57	12.0		
		327.0	65	15.2		
45	1941 II	333.9	60	2.0	+0.780	+1.06 ±0.25
	1941 IV	429.8	77	10.0		
	1943 V	439.5	91	14.8		
50	1941 III	406.5	59	4.4	+0.946	+1.562±0.082
	1941 IV	408.8	80	10.0		
	1943 IV	411.8	83	15.2		

sea level:

$$\frac{1}{L}\frac{dL}{dt}=0.16\pm0.03 \text{ percent per °C.}$$

The temperature effect of the meson intensity at sea level,

$$\frac{1}{I_\mu}\frac{dI_\mu}{dt}=\frac{L}{l}\times 0.16 \text{ percent per °C,}$$

can theoretically be as well greater or smaller than 0.16 percent in conformance to the mean free path of the mesons before decay, l being smaller or greater than L. For the temperature effect of the neutrini, created at the decay of the mesons, we obtain:

$$\frac{1}{I_\nu}\frac{dI_\nu}{dt}=\frac{L}{l}\frac{\exp[-L/l]}{1-\exp[-L/l]}\cdot 0.16 \leqq 0.16 \text{ percent per °C;}$$

hence, theoretically, it can only have values smaller or equal to 0.16 percent per °C, whereas the experimental value is about five times greater.

[1] L. Jánossy and G. D. Rochester, Proc. Roy. Soc. London **A183**, 186 (1944).
[2] J. Barnóthy and M. Forró, Zeits. f. Physik **104**, 744 (1937); Phys. Rev. **53**, 848 (1938); Phys. Rev. **55**, 868 (1939); Phys. Rev. **58**, 844 (1940).
[3] J. Barnóthy, Zeits. f. Physik **115**, 140 (1940).

Atomic Absorption Coefficient for X-Rays

JOHN A. VICTOREEN
The Victoreen Instrument Company, Cleveland, Ohio
September 15, 1947

IN a previous paper[1] it was shown that the mass absorption coefficient for any given element could be calculated from the empirically derived expression:

$$\frac{\mu}{\rho}=(C\lambda^3-D\lambda^4)+\sigma_e Z\frac{N_0}{A}.$$

Further work has led to an expression which may be of theoretical significance.

It has been found that the atomic absorption coefficient τ_a, for any element and between any two critical absorption wave-lengths, may be represented within the accuracy of observational error by the expression:

$$\tau_a=\left\{\frac{\nu_1\nu_2}{\nu^3}-\frac{\nu_1\nu_2\nu_3}{\nu^4}\right\}K,$$

where ν is the frequency of the incident radiation, and ν_1, ν_2, and ν_3 are apparently critical frequencies resulting from rational quantized transitions within an atom of atomic number Z. K is an invariable constant. A manuscript on the subject is now in preparation and will be published shortly.

[1] J. A. Victoreen, J. App. Phys. **14**, 95 (1943).

Electrostrictive Effect in Barium Titanate

W. P. MASON
Bell Telephone Laboratories, Murray Hill, New Jersey
September 10, 1947

IN a recent paper[1] and in his thesis Shepard Roberts has demonstrated a new type of electrostrictive effect in a ceramic piece made up of polycrystalline barium bitanate. In this effect mechanical resonances can be excited by a small applied alternating voltage when either a high direct-current field, or a remanence polarization induced by a high direct-current field, are present. Roberts identified the lowest frequency modes as radial vibrations of the disk, but, although his data show it, he failed to identify the thickness mode with a very high electromechanical coupling. Further measurements have been made by the writer of the radial and thickness modes of barium titanate disks as a function of the applied field. Considering the device as an electromechanical vibrator,[2] the data on the resonance frequency, the separation of resonance and antiresonance frequencies, and the capacity of the crystal give enough data to determine the electromechanical coupling factor k, the elastic constant for radial vibrations (Young's modulus), the elastic constant for the longitudinal thickness mode ($c_{11}=\lambda+2\mu$), and the electrostrictive constants. The elastic constants are

$$Y_0=\frac{\mu(3\lambda+3\mu)}{\lambda+\mu}=9.1\times10^{11} \text{ dyne/cm}^2; \quad (1)$$
$$\lambda+2\mu=1.16\times10^{12} \text{ dyne/cm}^2.$$

From these we can determine

$$\lambda=4.4\times10^{11};\quad \mu=3.8\times10^{11};$$
$$\sigma=\text{Poisson's ratio}=0.27. \quad (2)$$

The electrostriction constants for the two modes are shown plotted in Fig. 1 for ascending and descending

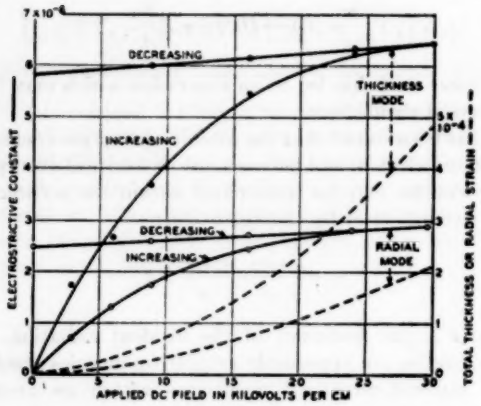


FIG. 1.

voltage gradients. The thickness constant is about twice that for the radial motion. After removing the polarization voltages of 30,000 volts per centimeters, the remanence polarization is enough to give a high coupling. Since the electrostriction constant is the slope of the electrostriction expansion-voltage curve, the total thickness and radial expansions can be obtained by integrating these curves, and the results are shown by the dotted lines of Fig. 1. The total increase in thickness is 5 parts in 10,000 up to a voltage gradient of 30,000 volts per centimeter.

The discovery of the thickness effect, and the ratio of its electrostrictive constant of about 2 to 1 to that for the radial mode, allows one to obtain a mechanism for this effect. Barium titanate above 120°C has a cubic structure, but at room temperature it becomes ferro-electric and has expanded by two-thirds of one percent[3] along the direction of the ferro-electric axis, and contracted one-third of one percent along the other two axes. In a polycrystalline body, the domains are distributed in all directions. The application of a high direct-current field lines up a certain percentage of the domains along the direction of the field. Since if all the domains were lined up the thickness would increase by two-thirds percent or 6.6×10^{-3}, the data of Fig. 1 indicate that

$$\frac{5\times10^{-4}}{6.6\times10^{-3}}=.075, \tag{3}$$

or 7.5 percent of the domains are lined up by a field of 30,000 volts per centimeter. Since the small alternating voltage superimposed on the direct-current field is too small to cause domains as a whole to change orientation, its principal effect is probably the transferring of molecules from domains not lined up with the field to domains lined up with the field, across their common boundary, when the alternating current field adds to the direct-current field and the reverse process when the two are opposing. When more molecules are in the domains lined up with the field, the crystal increases in thickness by a certain percentage and decreases in a radial direction by about half as much. When fewer molecules are in the lined-up domains, the reverse process occurs. Since the transfer of molecules from one domain to another lags behind the applied field, this is the cause of the dielectric hysteresis that lowers the Q of the vibrations to about 15. The transfer of molecules between adjacent domains can go on for an alternating field without a biasing polarization, but it cannot change the thickness unless a polarization bias exists.

[1] Shepard Roberts, "*Dielectric and piezoelectric properties of barium titanate*," Phys. Rev. **71**, 890 (1947).
[2] See, for example, W. P. Mason, "*Properties of monoclinic crystals*," Phys. Rev. **70**, 705 (1946).
[3] H. D. Megaw, Proc. Roy. Soc. **189**, 261 (1947).

Angular Correlation of Successive Gamma-Ray Quanta

EDWARD L. BRADY AND MARTIN DEUTSCH
Massachusetts Institute of Technology, Cambridge, Massachusetts
September 10, 1947

THEORETICAL considerations[1,2] predict a directional correlation of successive quanta of the form

$$W(\theta)=1+\sum_{i=1}^{l} A_i\cos^{2i}\theta$$

if $2l$ is the highest multipole order occurring. Attempts to demonstrate this effect experimentally have heretofore been inconclusive.[3] We have studied coincidences between successive gamma-rays of Co^{60}, Sc^{46}, Y^{86} (106 day), and Cs^{134} at angles 180° and 90° between the counters, and found a pronounced anisotropy in the first two named and no correlation within the experimental error in the last two. Our results, together with the gamma-ray energies concerned, are shown in Table I. The quantity $(W(\pi)$

TABLE I. Anisotropy of gamma-ray coincidences.

Source	Co^{60}	Sc^{46}	Y^{86}	Cs^{134}
ϵ	0.21 ± 0.025	0.20 ± 0.035	-0.05 ± 0.03	0.01 ± 0.04
Gamma-rays	1.1, 1.3	0.89, 1.12	0.91, 1.89	0.58, 0.78
Mev				
Reference	4	5	6	7

$-W(\pi/2))/W(\pi/2)$ should be equal to ΣA_i. Our results for this quantity, denoted by ϵ in Table I, should be increased by about six percent, because of finite angular resolution, and by an unknown, but probably small, correction arising from scattering in the source. The angular resolution of about ±15° was determined experimentally by measurements on annihilation radiation. Possible important effects attributable to interatomic fields[1] were shown to be absent in the case of Co^{60}, which was studied in the form of solid chloride, in aqueous solution, and as gaseous cobalt-nitrosyl-carbonyl. The respective values of ϵ were 0.19 ± 0.04, 0.20 ± 0.04, and 0.22 ± 0.04.

Spurious angular correlations can be caused by small contaminations by positron-emitting substances. This was ruled out in the case of our samples of cobalt and scandium by observing, with greatly improved angular resolution, coincidences from stronger samples of the same preparations. In this arrangement the effect of annihilation radiation would be greatly emphasized. It was found that no

observable part of our results was due to positron contamination.

An apparent anisotropy can be caused by secondary quanta scattered from one counter to the other. This effect was observed and appropriate lead shielding arranged to remove it. Its absence in the final experiments was shown by observations using Fe^{59}, which emits[8] two noncoincident gamma-rays of 1.1 Mev and 1.3 Mev. The coincidence rate was equal to the expected chance and cosmic-ray background, and was independent of angle.

It is interesting to note that the observed values of ϵ are consistent, within the experimental errors, with the values of ΣA_i predicted,[1] if the gamma-rays of Co^{60} and Sc^{46} are quadrupole radiation and the angular momenta of the nuclear states are 4, 2 and 0 in order of decreasing excitation energy in both cases. These spin assignments are in good agreement with considerations based on selection rules.[4] Further experiments are in progress. We wish to thank Mr. A. C. Miller for help in taking data. This work was supported in part by the Office of Naval Research.

[1] D. R. Hamilton, Phys. Rev. **58**, 122 (1940).
[2] G. Goertzel, Phys. Rev. **70**, 897 (1946).
[3] Kikuchi, Watase, and Itoh, Zeits. f. Physik **119**, 185 (1942); R. Beringer, Phys. Rev. **63**, 23 (1943); W. M. Good, Phys. Rev. **70**, 978 (1946).
[4] M. Deutsch, L. G. Elliott, and A. Roberts, Phys. Rev. **68**, 193 (1945).
[5] A. Miller and M. Deutsch, Bull. Am. Phys. Soc. (1947) (Montreal meeting).
[6] J. R. Downing, M. Deutsch, and A. Roberts, Phys. Rev. **60**, 470 (1941).
[7] K. Siegbahn and M. Deutsch, Phys. Rev. **71**, 483A (1947).
[8] M. Deutsch, J. R. Downing, L. G. Elliott, J. W. Irvine, Jr., and A. Roberts, Phys. Rev. **62**, 3 (1942).

Preliminary Analysis of the Microwave Spectrum of SO_2*

B. P. Dailey, S. Golden,** and E. Bright Wilson, Jr.
Mallinckrodt Chemical Laboratory, Harvard University, Cambridge, Massachusetts

September 10, 1947

THE purpose of the present note is to indicate the progress made to date on the analysis of the rotational spectrum of SO_2 as observed in the frequency range 20,000–30,000 mc/sec.[1,2]

Although thirty-two lines have been observed, approximate measurements of intensity and of temperature coefficients of intensity indicate that only five lines correspond to transitions of moderately low J of $S^{32}O_2^{16}$ in the ground vibrational state. The remainder includes lines presumably caused by high J transitions, by transitions of molecules in higher vibrational states, and by isotopic species such as $S^{34}O_2^{16}$, etc. In the analysis, SO_2 was considered to be a rigid asymmetric rotor, with effective moments of inertia replacing equilibrium moments of inertia.

From the Stark splitting of the lines at 23,413 and 29,460 mc/sec., it was possible to determine the values of the main quantum numbers, J, involved in the transitions. Without placing some limits on the range of values of the moments of inertia it seems quite difficult to pursue the analysis beyond this point. If however, the S–O distance is limited to values between the single-bond and triple-bond distances determined from normal covalent radii (1.37A–1.70A), and the O–S–O angle is limited to 70–180°, a unique identification of the two resolvable lines can be made.

These transitions permitted a sufficiently accurate determination of the rotational constants to predict a line that could be easily identified with the line observed at 25,392 mc/sec. Using these three transitions (of lowest J), a readjustment of the rotational constants was made that permitted additional lines to be identified with little difficulty. After the identification was made, the rotational constants were evaluated by applying the method of least squares to the five lines listed in Table I.

Consistency of the identification was tested in two ways:

(1) Comparison of the observed quantity

$$\Delta^0 = I_c^0 - I_a^0 - I_b^0$$

with that calculated from the fundamental frequencies of $S^{32}O_2^{16}$ with the aid of the formula given by Darling and Dennison.[3] The results are

$$\Delta^0(\text{obs.}) = 0.20\times10^{-40}\ \text{g cm}^2,$$
$$\Delta^0(\text{calc.}) = 0.24\times10^{-40}\ \text{g cm}^2.$$

(2) Comparison of the observed Stark effect with that computed from the rotational constants. The results for the component frequencies measured relative to the unsplit line are

$[6_{1,5}\rightarrow5_{2,4}]$ at 23,412 mc/sec.
$\Delta\nu(\text{obs.}) = (0.27-0.013M^2)\epsilon^2(\text{e.s.u.})$,
$\Delta\nu(\text{calc.}) = (0.224-0.0110M^2)\epsilon^2(\text{e.s.u.})$,
$[3_{1,3}\rightarrow4_{0,4}]$ at 29,460 mc/sec.
$\Delta\nu(\text{obs.}) = (0.31-0.034M^2)\epsilon^2(\text{e.s.u.})$,
$\Delta\nu(\text{calc.}) = (0.294-0.0325M^2)\epsilon^2(\text{e.s.u.})$.

The effective moments of inertia based upon the above assignment are

$$I_c^0 = 95.14\times10^{-40}\ \text{g cm}^2,$$
$$I_b^0 = 81.16\times10^{-40}\ \text{g cm}^2,$$
$$I_a^0 = 13.78\times10^{-40}\ \text{g cm}^2.$$

Since the effect of centrifugal distortion was neglected, the assignments given and the values of the effective moments of inertia must be taken to be tentative. Because of the influence of zero-point vibration and other factors, it is not possible to determine the molecular parameters directly from the effective moments of inertia. However,

TABLE I. Tentatively identified lines of the rotational spectrum of SO_2.

ν, mc/sec.	Tentative assignment	Relative intensity	$\frac{I_{-70°C}}{I_{27°C}}$	Comments
20,420	$13_{2,12}\rightarrow12_{3,9}$	2	4	
23,413	$6_{1,6}\rightarrow5_{2,4}$	1.5	5	5 components observed
24,037 or 24,083	$9_{1,9}\rightarrow8_{2,6}$	5	2	
		5	3	
25,392	$7_{2,6}\rightarrow8_{1,7}$	10	3	2 components observed
29,460	$3_{1,3}\rightarrow4_{0,4}$	—	—	2 components observed

an estimate may be made using the values obtained for the smallest and intermediate moments of inertia. The results are

effective S−O distance =1.433A,
effective O−S−O angle=119.5°.

These are in good agreement with the results obtained from electron-diffraction measurements.

* The research reported in this document was made possible through support extended Harvard University by the Navy Department (Office of Naval Research) under Office of Naval Research Contract N5ori-76.
** National Research Council Predoctoral Fellow.
[1] R. H. Hughes and E. B. Wilson, Jr., Phys. Rev. **71**, 562 (1947).
[2] B. P. Dailey, Phys. Rev. **72**, 84 (1947); B. P. Dailey and E. B. Wilson, Jr., paper presented at regional meeting of the American Physical Society, Amherst, Massachusetts, May 31, 1947.
[3] B. T. Darling and D. M. Dennison, Phys. Rev. **57**, 128 (1940).

Excitation Curves for the Reactions $C^{12}(d, dn)C^{11}$ and $C^{12}(\alpha, \alpha n)C^{11}$ at High Energies

R. L. Thornton and Richard W. Senseman
Radiation Laboratory, University of California, Berkeley, California
September 22, 1947

THE energy dependence of the reactions $C^{12}(d, dn)C^{11}$ and $C^{12}(\alpha, \alpha n)C^{11}$ has been investigated to energies of 195 and 390 Mev, respectively. The usual technique was employed of bombarding a stack of plates of pure graphite of total thickness sufficient to contain the range of the deuterons or alpha-particles and determining the induced C^{11} activity (20.5 min.) of the individual plates. The ion beam of the 184-inch cyclotron was deflected inward by a pulsed electrostatic deflector of angular length about 120 degrees. The deflected beam reaches its maximum radius 360 degrees from the start of the deflector and there has an increase in radius of two to three inches.[1] The deflected beam is incident on the stack of plates which was mounted in a lead-shielded box on the end of the target probe and could be withdrawn through a vacuum lock. The ion beam was defined by a window in a lead plate one-inch thick immediately in front of the graphite plates. Experiment showed that the overlap of the plates over the window opening was sufficient to contain the beam within the area of the plates. Plates of both $\frac{1}{4}$- and $\frac{1}{16}$-inch thickness were used simultaneously in determining the excitation curve. The majority of the thin plates were used in the region where the curve was changing most rapidly, but others were interspersed between $\frac{1}{4}$-inch plates at other points. Since for the same current and energy the thin plates gave lower activity, it was necessary to adjust the two resulting curves. This was done by multiplying the measured activities of the thin plates by a factor which would bring the curves into coincidence at the high energy region where the excitation curve was found to be slowly varying.

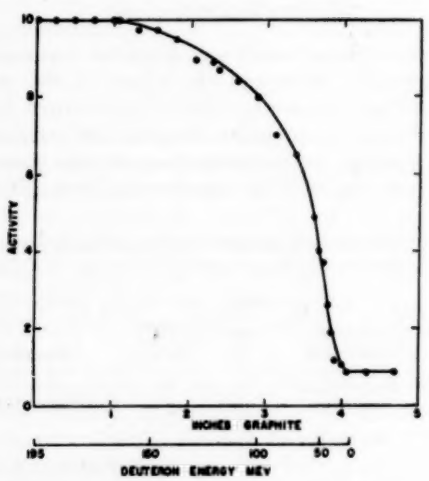


FIG. 1. Relative yield of C^{11} activity as a function of deuteron range in graphite.

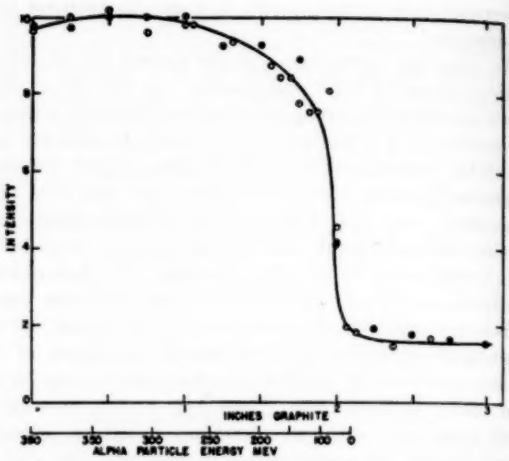


FIG. 2. Relative yield of C^{11} activity as a function of alpha-particle range in graphite. The open and solid circles give results from independent runs adjusted to average coincidence at high energies.

The radioactive decay curves of the plates show an exponential decay of period corresponding to C^{11}. In the case of deuteron activation, an indication was observed of a shorter period on plates near the end of the deuteron range. This is presumably caused by the reaction $C^{12}(d, n)N^{13}$ which, however, was not observed at higher energies. The results obtained are presented in Figs. 1 and 2. It will be noted that residual activity is observed at thicknesses greater than the range of the ions. This is due to fast neutron activation by the $(n, 2n)$ reaction; in the alpha-particle curve additional background arises from the deuteron contamination of the ion beam.

It is interesting to note that in neither curve is there apparent a tendency for the yield to fall off at high energies because of the increasing probability of more complicated competing reactions. This would suggest that the maximum energy of the incident particle is not, in general, available for such reactions, but that it is more probable that only a fraction of this energy is transferred to the nucleus. On the other hand, experiments with elements of higher atomic number, and the frequent

production of stars, show clearly that the probability of the emission of many nucleons is significant. Further experiments are in progress with other reactions which it is hoped will clarify the mechanism of nuclear reactions in this energy region.

This work was supported by the Atomic Energy Commission under Contract W-7405-Eng-48 with the University of California.

[1] This deflector system will be described in detail in a forthcoming paper by Dr. Wilson M. Powell.

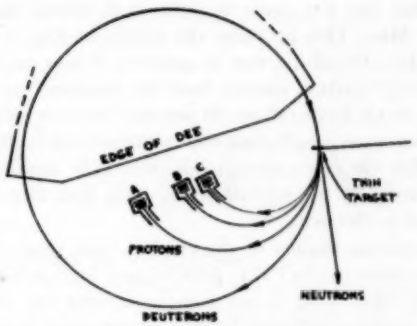


FIG. 1. Schematic diagram of experimental set-up in the 184-in. cyclotron. *A*, *B*, and *C* indicate the three positions of the houses containing the carbon plates. These are below the level of the deuteron orbits.

Erratum: Spectral Location of the Absorption Due to Color Centers in Alkali Halide Crystals

[Phys. Rev. **72**, 341 (1947)]

HENRY F. IVEY
Research Laboratory, Westinghouse Electric Corporation, Bloomfield, New Jersey

REFERENCE (6) should read: E. Mollwo, as reported by Pohl in reference (*g*). It now appears as ". . . reference (3)."

In the table, the reference for the value 5700 for the R_2-band of KF should be (*d*). It now appears as (6).

In the table, the reference for the value 7200 for the *F*-band of RbBr should be (*f*). It now appears as (8).

In the heading of the second column of the table, "Inter-Ionic Salt Distance," the word "Salt" should be removed. This word should be the heading of the first column, but its omission is not confusing.

Excitation Curve for the Reaction $C^{12}(p, pn)C^{11}$ up to 140 Mev

WARREN W. CHUPP AND EDWIN M. MCMILLAN
Radiation Laboratory, University of California, Berkeley, California

September 22, 1947

THE work of Helmholz, McMillan, and Sewell[1] and Serber[2] has shown that neutrons are stripped off of deuterons when they strike an internal cyclotron target; one naturally expects a similar process in which protons are set free. These protons, having a mean energy half that of the deuterons, will move in circles passing near the center of the cyclotron and with the expected energy spread resolved as in a mass spectrograph. Their distribution is being studied by Chupp, Gardner, and Taylor,[3] and the work reported in this letter shows that their intensity is quite adequate for nuclear experimental work. Stacks of carbon plates enclosed in copper houses, and with suitable defining slits in front, were placed as shown in Fig. 1, the houses being below the level of the circulating deuterons and tilted upward at a proper angle. After exposure to the proton beam, the carbon plates were taken out and their activities (20.5-min. C^{11}) measured on a G-M counter; the plot of the activities against position in the stack gives directly the excitation curve on a range scale.

A series of four runs was made with the plates in position *C*, where the incident proton energy was found to be 65 Mev. The carbon plates were disks $1\frac{1}{16}$-in. diameter by $\frac{1}{32}$ in. thick, with a surface density of 137 mg/cm². In order to establish the end of the range, plates of compressed boric acid having the same stopping power were interposed between some of the carbon plates; the resulting activity formed by $B^{11}(p, n)C^{11}$ with a low threshold gives a fiducial mark from which the range can be found within the resolving power of this experiment (about $\pm\frac{1}{32}$" from the energy spread admitted by the slit system). In Fig. 2, the points corresponding to energies below 65 Mev come from this series. A small neutron background of about 3 percent was subtracted from the carbon activities. Each point is a mean of all the measurements, and the spread found in individual values is of the same order as the irregularities in the curve.

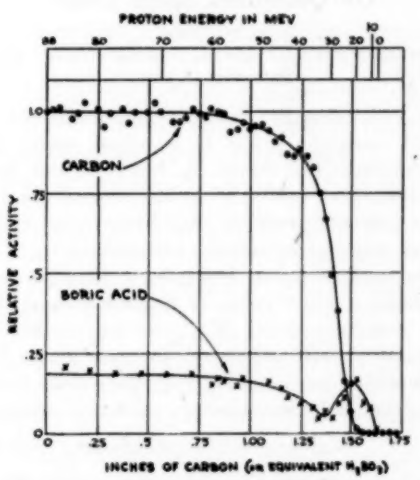


FIG. 2. Excitation curve for the C^{11} activity induced in carbon plates by protons entering at the left. Each point is plotted at a position corresponding to the surface of the plate on which the activity was measured. The lower curve represents the C^{11} activity found in H_3BO_3 plates between some of the carbon plates, and was used to find the end of the range. Points above 65 Mev come from one run and those below 65 Mev from four runs averaged together; these were adjusted in height to fit at 65 Mev. Another run extending the curve to 140 Mev is not included, since it indicates a constant cross section. One inch of carbon corresponds to 4.38 g/cm².

Another run was made in position B, where the energy was 88 Mev. This provides the points in Fig. 2 from 88 to 65 Mev. Finally, a run in position A was made, using $2''\times3''\times\frac{1}{4}''$ carbon plates; here the neutron background was relatively large (about 40 percent) because the number of protons was small, but the accuracy was sufficient to show that the cross section does not vary appreciably in going from 88 up to 140 Mev. This last run was not included in the plot.

The striking feature of this result is that there is little if any variation in the $C^{12}(p, pn)C^{11}$ cross section between 60 and 140 Mev. This is not consistent with the picture of proton capture followed by pn evaporation, because the competition of other reactions at high excitation energies would reduce the cross section rapidly. One must assume a non-capture excitation or a (p, n) exchange with roughly constant energy transfer, followed by evaporation of a single neutron or proton. Dr. Serber has pointed out[4] that this is a reasonable assumption in this energy range.

The activity in the boric acid is itself of some interest. The (p, n) reaction in B^{11} gives a peak near the end of the range, as expected for that type of reaction; the second rise at higher energy is attributed to reactions forming C^{11} from oxygen, since its magnitude is consistent with the activity observed in BeO traversed by the same proton current.

This work was done under the auspices of the Atomic Energy Commission, under Contract No. W-7405-Eng-48.

[1] A. C. Helmholz, E. M. McMillan, and D. Sewell, Phys. Rev., to be published.
[2] R. Serber, Phys. Rev., to be published.
[3] W. W. Chupp, E. Gardner, and T. B. Taylor, to be published.
[4] R. Serber, to be published.

On Quantized Space-Time

C. N. YANG
Department of Physics, University of Chicago, Chicago, Illinois
September 15, 1947

RECENTLY Snyder[1] has developed a theory of quantized space-time that is invariant under Lorentz transformations. The theory is, however, not invariant under translations. Indeed, as he has pointed out, to make the theory invariant under translations one must have space-time coordinates forming a continuum.

It does seem desirable, however, to have an invariant theory under a wider group of transformations than the Lorentz transformations. (E.g., we may want to have homogeneity as well as isotropy in space-time.) This can be accomplished by proposing that space-time is curved.

As an example let us consider a de Sitter universe

$$-x_0^2+x_1^2+x_2^2+x_3^2+\xi^2=R^2, \tag{1}$$

which is a pseudosphere in a five-dimensional flat space: x_0, x_1, x_2, x_3, ξ. The group of linear transformations that leaves the quadratic form $-x_0^2+x_1^2+x_2^2+x_3^2+\xi^2$ unchanged could be approximated by the product of the ordinary Lorentz group and the group of translations, for points on the pseudosphere in the region $|x_0|\ll R$, $|x_i|\ll R$. If we define $L_i/\hbar$, $M_i/\hbar$, and $Rp_0/\hbar$ and $Rp_i/\hbar$ as the nuclei of these transformations, we would get the angular- and linear-momenta operators in the de Sitter universe. The p's are given by

$$p_0=i\hbar\frac{\xi}{R}\frac{\partial}{\partial x_0}+i\hbar\frac{x_0}{R}\frac{\partial}{\partial\xi},$$
$$p_i=-i\hbar\frac{\xi}{R}\frac{\partial}{\partial x_i}+i\hbar\frac{x_i}{R}\frac{\partial}{\partial\xi}. \tag{2}$$

They satisfy

$$L_i=\xi(x_jp_k-x_kp_j)/R. \tag{3}$$

The commutators involving the p's are

$$[p_i, p_j]=i\hbar L_k/R^2,$$
and
$$[x_i, p_j]=i\hbar\xi\delta_{ij}/R. \tag{4}$$

Those involving the L_i, M_i, and x_μ's are as usual.

The above is, however, not the only theory that is invariant under the Lorentz transformations in five dimensions. We shall propose, in general, that x_0, x_1, x_2, x_3, ξ be Hermitian operators, and that after a Lorentz transformation in five dimensions they can be brought back to their original forms by a unitary transformation. This means that the nuclei of the unitary transformations, which we define to be $L_i/\hbar$, $M_i/\hbar$, and $Rp_0/\hbar$, and $Rp_i/\hbar$, still satisfy the same commutation relations (4). The commutation relations between the operators x_0, x_1, x_2, x_3, ξ are not fixed by the requirement of invariance. If we propose that they are given as follows,

$$\begin{aligned}[x_i, x_j]&=ia^2L_k/\hbar, & [x_0, x_i]&=ia^2M_i/\hbar,\\ [\xi, x_i]&=ia^2Rp_i/\hbar, & [\xi, x_0]&=ia^2Rp_0/\hbar,\end{aligned} \tag{5}$$

it is evident that the 15 operators $L_i/\hbar$, $M_i/\hbar$, $Rp_\mu/\hbar$, x_μ/a and ξ/a satisfy the same commutation relations as the nuclei of the group of Lorentz transformations in six dimensions with the basic quadratic form $-\eta_0^2+\eta_1^2+\eta_2^2+\eta_3^2+\eta^2+\zeta^2$. So a possible solution is

$$\begin{aligned}L_i&=i\hbar\left(\eta_k\frac{\partial}{\partial\eta_j}-\eta_j\frac{\partial}{\partial\eta_k}\right),\\ M_i&=i\hbar\left(\eta_0\frac{\partial}{\partial\eta_i}+\eta_i\frac{\partial}{\partial\eta_0}\right),\\ p_0&=\frac{i\hbar}{R}\left(\eta\frac{\partial}{\partial\eta_0}+\eta_0\frac{\partial}{\partial\eta}\right), & p_i&=\frac{i\hbar}{R}\left(\eta\frac{\partial}{\partial\eta_i}-\eta_i\frac{\partial}{\partial\eta}\right),\\ x_0&=ia\left(\zeta\frac{\partial}{\partial\eta_0}+\eta_0\frac{\partial}{\partial\zeta}\right), & x_i&=ia\left(\zeta\frac{\partial}{\partial\eta_i}-\eta_i\frac{\partial}{\partial\zeta}\right),\end{aligned} \tag{6}$$

and

$$\xi=ia\left(-\zeta\frac{\partial}{\partial\eta}+\eta\frac{\partial}{\partial\zeta}\right).$$

For this solution the eigenvalues of the space coordinates are discrete.

Returning to the general case, we should add Eq. (1) to the general Eqs. (4) and (5).

If we put $R=\infty$ we would get the special solution (2).

[1] H. Snyder, Phys. Rev. 71, 38 (1947).